

## THESIS / THÈSE

### MASTER EN SCIENCES MATHÉMATIQUES

#### Quelques algorithmes pour la minimisation de normes polyédrale et euclidienne

Bihain, André

*Award date:*  
1978

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX  
NAMUR

ANNEE ACADEMIQUE : 1977 - 1978

QUELQUES ALGORITHMES POUR  
LA MINIMISATION DE NORMES  
POLYEDRALE ET EUCLIDIENNE

NGUYEN VAN KIEN

MEMOIRE PRESENTE POUR L'OBTENTION DU GRADE  
DE LICENCE EN SCIENCES MATHÉMATIQUES  
PAR

PROMOTEUR :

V.H. NGUYEN.

André BIHAIN



Je remercie vivement  
Monsieur V.H. NGUYEN pour les  
conseils judicieux qu'il m'a  
donnés lors de la réalisation  
de ce mémoire.

## TABLE DES MATIERES

INTRODUCTION .....	p. I.1.
CHAPITRE I : Normes Polyédrales et un Algorithme d'Echange Pour des Problèmes d'Appro- ximation Linéaire .....	1.1.
§ 1. Normes Polyédrales - Définitions .....	1.1.
§ 2. Approximation Discrète .....	1.10.
2.1. Définition	
2.2. Caractérisation d'une Solution Optimale du Problème (2.1)	
§ 3. Généralisation aux Problèmes avec Contraintes	1.18.
§ 4. Un Algorithme d'Echange .....	1.21.
4.1. Notion Intuitive	
4.2. Algorithme - Description	
§ 5. Convergence de l'Algorithme .....	1.41.
Conclusions .....	1.44.
Références .....	1.45.

CHAPITRE II : Un Algorithme de Descente Pour une Minimisation en Norme Polyédrale .....	p. 2.1.
§ 1. Introduction et Position du Problème .....	2.1.
§ 2. Caractérisation d'un Point Stationnaire ....	2.5.
§ 3. Un Algorithme de Descente .....	2.9.
§ 4. Hypothèses Suffisantes de Convergence .....	2.22.
Conclusions .....	2.34.
Références .....	2.37.
 CHAPITRE III : Un Algorithme de Type Levenberg Pour une Minimisation en Norme Polyédrale .....	 3.1.
§ 1. Introduction .....	3.1.
§ 2. Bases Pour le Développement d'un Algorithme .	3.3.
§ 3. Algorithme .....	3.10.
3.1. Idée Intuitive	
3.2. Convergence	
3.3. Forme de l'Algorithme	
§ 4. Vitesse de Convergence de l'Algorithme .....	3.21.
Références .....	3.25.

CHAPITRE IV : Un Algorithme de Levenberg Pour la Résolution d'un Problème de Minimisation en Norme Euclidienne .....

p. 4.1.

§ 1. Introduction ..... 4.1.

§ 2. Quelques Rappels sur les Moindres Carrés .... 4.3.

§ 3. Algorithme Théorique ..... 4.5.

3.1. Idée et Hypothèse de Travail

3.2. Convergence

3.3. Algorithme

§ 4. Remarques Complémentaires ..... 4.21.

Conclusions ..... 4.26.

Références ..... 4.27.

CHAPITRE V : Implémentation et Résultats Numériques ... 5.1.

§ 1. Normes Polyédrales ..... 5.1.

5.1. Problèmes Linéaires

5.2. Problèmes Non-Linéaires

§ 2. Norme Euclidienne ..... 5.16.

ANNEXE I : Eléments de Programmation Linéaire .....	p. A1.1.
§ 1. Dual d'un Problème Linéaire .....	A1.1.
§ 2. Solution Optimale d'un Problème Linéaire ....	A1.3.
§ 3. Solution Optimale Duale .....	A1.6.
§ 4. Théorème Faible des Ecartés Complémentaires ..	A1.10.
ANNEXE II : : Sous-Routines FORTRAN .....	A2.1.

## INTRODUCTION

Parmi les problèmes d'optimisation, les problèmes où la fonction-objectif se présente sous la forme d'une norme interviennent très souvent en approximation et en statistique.

La littérature concernant ce problème

$$\min \|f(x)\|$$

où  $f$  est une fonction vectorielle, est abondante. Malheureusement, celle-ci concerne presque exclusivement la norme euclidienne. La raison principale en est que la dérivée de cette norme s'exprime d'une manière simple et pratique.

Le mémoire aborde ce problème mais développe surtout une théorie où la norme utilisée est une norme dite *polyédrale* et englobant, d'une manière assez élégante, les normes  $l_1$  et  $l_\infty$ .

Cette théorie a été proposée par D.H. ANDERSON et M.R. OSBORNE dans trois articles parus dans la revue "Numerische Mathematik" en 1976 (No 26) et 1977 (No 28).

Nous basant sur ces textes, nous analysons, dans un premier chapitre, un algorithme d'échange s'appliquant au cas linéaire, c'est-à-dire

$$\min_{h \in \mathbb{R}^p} \|f + A h\|_B$$

où  $A$  et  $f$  sont respectivement une matrice et un vecteur donnés.  $\|\cdot\|_B$  désigne la norme polyédrale dont la définition précise est donnée en début de ce chapitre. Nous montrons également que cet algorithme d'échange généralise en fait l'algorithme d'échange classique dans le cas de la norme  $l_\infty$ .



Le second chapitre utilise cet algorithme afin de résoudre le problème de minimisation non-linéaire

$$\min_{x \in \mathbb{R}^p} \|f(x)\|_B,$$

par une méthode de type Gauss-Newton.

Le troisième chapitre développe un algorithme de type Levenberg en norme polyédrale qui apporte une solution aux problèmes que la méthode de Gauss-Newton ne peut pas traiter, c'est-à-dire les problèmes où la matrice jacobienne n'est pas de rang maximal.

Le sujet, dans le quatrième chapitre, est étudié en norme euclidienne. Analysant un article "Nonlinear Least Squares - The Levenberg Algorithm Revised" de M.R. OSBORNE (à paraître), nous donnons une nouvelle version de l'algorithme de Levenberg qui, par le choix du paramètre de descente, en assure la convergence globale.

Le dernier chapitre résume les résultats numériques obtenus lors de l'implémentation des algorithmes des chapitres I, II et IV.

Une première annexe explicite les résultats de programmation linéaire que nous utilisons tandis qu'une seconde annexe rassemble les programmes FORTRAN implémentés.



## CHAPITRE I :

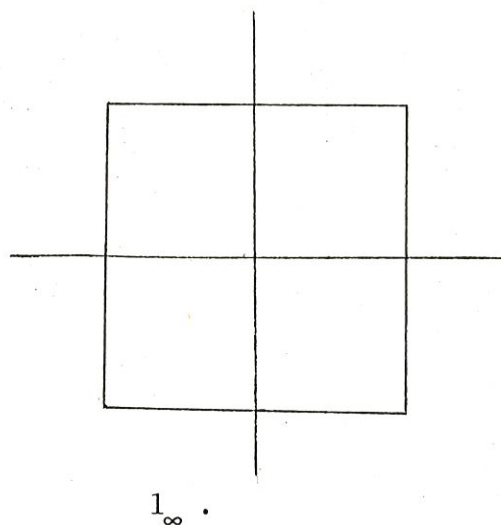
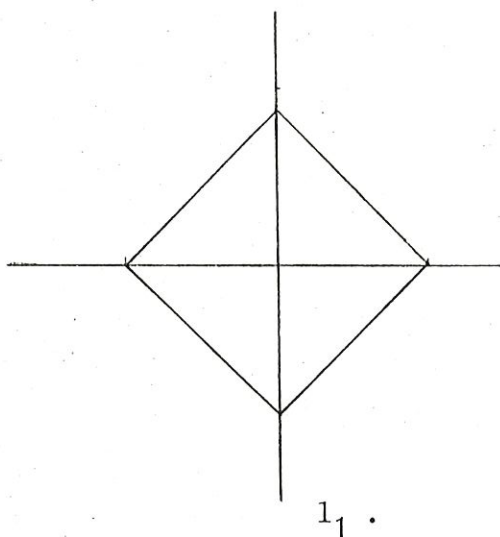
NORMES POLYÉDRALES ET UN ALGORITHME D'ÉCHANGE  
POUR DES PROBLÈMES D'APPROXIMATION LINÉAIRE

## § 1. NORMES POLYEDRALES - DEFINITIONS

---

Le concept de norme polyédrale correspond, comme nous allons le voir, d'une part à une généralisation naturelle des normes  $l_1$  et  $l_\infty$  et rejoint d'autre part la théorie des ensembles et fonctions convexes.

Afin de rassembler ces deux *notions*, rappelons-nous les structures semblables des boules unités relatives aux normes  $l_1$  et  $l_\infty$  dans  $\mathbb{R}^2$  :



Cette structure de boule délimitée par un ensemble de droites dans  $\mathbb{R}^2$ , ou d'hyperplan dans  $\mathbb{R}^n$ , peut se généraliser en utilisant les résultats de l'analyse convexe.

Avant de résumer ces résultats, rappelons quelques définitions.

### Définitions.

---

- Si  $f : S \rightarrow \mathbb{R}$  ( $S \subset \mathbb{R}^n$ ), on appelle *épigraphe* de  $f$  l'ensemble  $\{(x, \mu) ; x \in S, \mu \geq f(x)\} \subset \mathbb{R}^n \times \mathbb{R}$ .

- Si  $E$  est un espace vectoriel sur  $\mathbb{C}$ , un sous-ensemble  $A \subset E$  est dit *équilibré* ssi, pour tout  $x \in A$ , pour tout  $\lambda \in \mathbb{R}$ ,

$$|\lambda| \leq 1, \quad \lambda x \in A.$$

- Si  $E$  est un espace vectoriel sur  $\mathbb{C}$ , un sous-ensemble  $A \subset E$  est dit *absorbant* ssi, pour tout  $x \in E$ , il existe  $\varepsilon > 0$  tel que

$$\lambda x \in A, \quad \forall \lambda \in \mathbb{R}, \quad |\lambda| \leq \varepsilon.$$

- Si  $E$  est un espace vectoriel sur  $\mathbb{C}$ , on appelle *semi-norme* sur  $E$  toute fonction  $p : E \rightarrow \mathbb{R}^+$  vérifiant les conditions suivantes :

- i)  $\forall x, y \in E, \quad p(x+y) \leq p(x) + p(y)$  (sous-additivité) ;
- ii)  $\forall \lambda \in \mathbb{C}, \quad p(\lambda x) = |\lambda| p(x)$  (homogénéité en valeur absolue) .

Rappelons encore un théorème connu d'analyse fonctionnelle dont la démonstration se trouve dans [1, pp. 146-147] .

### Théorème 1.1.

---

Soient  $E$  un espace vectoriel topologique et  $B$  un ensemble convexe, équilibré, absorbant et d'intérieur non-vide.

Il existe alors une unique semi-norme  $p$  sur  $E$  définie par

$$p(x) = \inf \{ \alpha > 0 ; x \in \alpha B \}$$

où

$$B = \{x ; p(x) \leq 1\}$$

Ces quelques rappels terminés, revenons à notre but, c'est-à-dire généraliser les normes  $l_1$  et  $l_\infty$  ou les normes dont les boules unités peuvent s'exprimer sous forme d'inégalités.

Ceci nous conduit à la définition suivante :

Soient  $B$  une matrice  $(m \times n)$  et  $b$  un vecteur  $(m \times 1)$  tels que :

$$\text{i) } \text{rg } B = n \quad ; \quad (1.1)$$

$$\text{ii) } b > 0 \text{ , c'est-à-dire } b_i > 0 \quad \forall i = 1, \dots, m \quad ; \quad (1.2)$$

$$\text{iii) } \text{l'ensemble } S_B = \{x \in \mathbb{R}^n ; Bx \leq b\} \text{ est borné} \quad ; \quad (1.3)$$

$$\text{iv) } \text{pour tout } X = (X_1, \dots, X_n)^T \in \mathbb{R}^n, \text{ on a}$$

$$B X \leq b \Leftrightarrow B |X| \leq b$$

où

$$|X| = (|X_1|, \dots, |X_n|)^T \quad (1.4)$$

(l'indice supérieur  $T$  désigne la transposée);

nous appellerons *norme polyédrale* toute norme ayant  $S_B$  comme boule unité.

Vérifions l'existence d'une telle norme et, pour cela, utilisons le Théorème 1.1.  $S_B$  est immédiatement un ensemble convexe d'intérieur non-vide. Il est absorbant car

$$\forall x \in \mathbb{R}^n, \exists \varepsilon = \min_{(Bx)_i \neq 0} \left\{ \frac{b_i}{|(Bx)_i|} \right\} \text{ tel que,}$$

pour tout  $\lambda$ ,  $|\lambda| < \varepsilon$ , on ait  $\lambda x \in S_B$ .

Il est immédiatement équilibré car

$$\begin{aligned} x \in S_B &\Leftrightarrow Bx \leq b \\ &\Leftrightarrow B|x| \leq b && \text{par (1.4)} \\ &\Leftrightarrow |\lambda| B|x| \leq b && \forall |\lambda| \leq 1 \\ &\Leftrightarrow B|\lambda x| \leq b && \forall |\lambda| \leq 1 \\ &\Leftrightarrow B\lambda x \leq b && \text{par (1.4)} \\ &\Leftrightarrow \lambda x \in S_B && \forall |\lambda| \leq 1 \end{aligned}$$

Dès lors, l'application du Théorème 1.1 permet d'affirmer qu'il existe une semi-norme  $p_B(x)$  définie par

$$p_B(x) = \inf \{ \alpha > 0 ; x \in \alpha S_B \} \quad (1.5)$$

et telle que

$$S_B = \{ x : p_B(x) \leq 1 \} \quad (1.6)$$

Pour terminer la justification de la définition, il nous faut voir que cette semi-norme est, en réalité, une norme, c'est-à-dire :

$$\forall x \in \mathbb{R}^n, x \neq \theta_n \text{ (origine de } \mathbb{R}^n) \Rightarrow p_B(x) \neq 0.$$

Preuve :

Soient  $\alpha > 0$  et  $x \in \mathbb{R}^n$  ;

$$x \in \alpha S_B \Leftrightarrow Bx \leq \alpha b \quad . \quad (1.7)$$

D'autre part,  $\text{rg } B = n$  , ce qui implique, que

$$x \neq \theta_n \Rightarrow Bx \neq \theta_m \quad .$$

Il existe donc  $i \in \{1, \dots, m\}$  tel que

$$(Bx)_i \neq 0 \quad .$$

Par la propriété (1.4) , nous pouvons supposer que

$$(Bx)_i > 0 \quad .$$

Dès lors, (1.7) est satisfaite si

$$\alpha \geq \frac{b_i}{(Bx)_i}$$

qui est donc une borne inférieure pour  $p_B(x)$  définie par (1.5) . $\square$

Commentaires.

---

(1) Le terme *norme polyédrale* se justifie certainement par le fait que les propriétés (1.1)-(1.4) donnent, comme boule unité de cette norme, un polyèdre convexe symétrique de  $\mathbb{R}^n$  .

Notons cependant que cette définition s'accorde avec la notion plus large de fonction polyédrale définie dans [2, pp. 170-173] , c'est-à-dire :



Une fonction convexe *polyédrale* est une fonction convexe dont l'épigraphe peut s'écrire sous la forme d'un ensemble fini d'inégalités :

$$\langle x, c_i \rangle \leq \beta_i \quad (i = 1, \dots, m)$$

où

$$\langle x, y \rangle = \sum_{j=1}^n x_j y_j$$

Il est immédiat que la norme définie ci-dessus vérifie cette définition.

- (2) La matrice  $B$  définissant la norme polyédrale possède une propriété très intéressante que nous utiliserons largement dans ce qui suit. Dès lors, nous préférons mettre en évidence cette caractéristique qui est une conséquence de la condition (1.4), c'est-à-dire :

Si la matrice  $B$  comprend la ligne  $(b_{i1}, \dots, b_{in})^T$ , elle comprend les lignes obtenues en gardant les nombres  $(|b_{i1}|, \dots, |b_{in}|)$  inchangés mais affectés des signes  $+$  ou  $-$  et ce, de toutes les manières possibles.

En effet, sans perdre de généralité, nous pouvons supposer que

$$b_i = 1, \quad \forall i \in \{1, \dots, m\}$$

De plus, le minimum, c'est-à-dire  $p_B(x) = \inf \{\alpha > 0, Bx \leq \alpha b\}$  est atteint. En effet, supposons ce minimum non atteint; donc, pour tout  $i = 1, \dots, m$ ,

$$(Bx)_i < p_B(x) b_i$$

Dès lors, en posant

$$-\epsilon = \min_i [(Bx)_i - p_B(x) b_i],$$

on obtient

$$(p_B(x) - \varepsilon) \in \{\alpha > 0, Bx \leq \alpha b\},$$

ce qui est impossible par définition de  $p_B(x)$ .

Dès lors, pour tout  $x \in \mathbb{R}^n$ , il existe  $q \in \{1, \dots, m\}$  tel que

$$p_B(x) = \sum_{j=1}^n B_{qj} x_j.$$

Changeons le signe d'une composante non-nulle de  $x$ , c'est-à-dire posons

$$x'_i = x_i \quad \text{si } i \neq i_0,$$

$$x'_i = -x_i \quad \text{si } i = i_0.$$

Dès lors, il existe  $q_0$  tel que

$$p_B(x) = \sum_{j=1}^n B_{q_0 j} x'_j.$$

Par la condition (1.4), on obtient

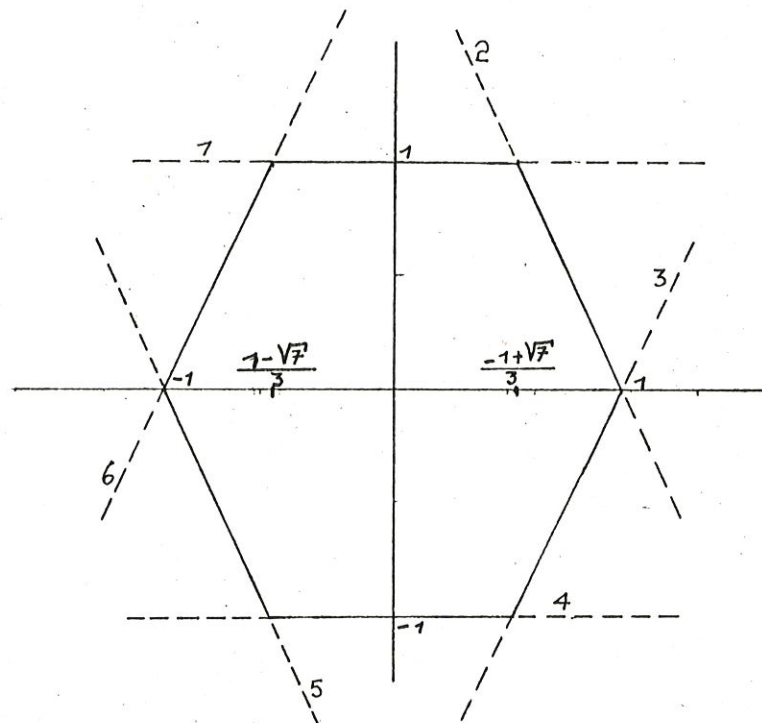
$$p_B(x) = p_B(x'),$$

ce qui veut dire que, à la ligne  $[B_{q_1}, \dots, B_{q_{i_0}}, \dots, B_{q_n}]^T$  correspond une ligne de la forme  $[B_{q_1}, \dots, -B_{q_{i_0}}, \dots, B_{q_n}]^T$ .  $\square$

Exemples.

- (1) Dans  $\mathbb{R}^2$ , déterminons la forme de la matrice  $B$  telle que la boule unité pour cette norme soit l'hexagone ci-après,





qui est simplement le polyèdre défini par les droites 1 à 6 .

On obtient, en écrivant l'équation de ces droites, le système d'inégalités suivant :

$$\begin{pmatrix} 1 & (4 - \sqrt{7}) / 3 \\ -1 & (-4 + \sqrt{7}) / 3 \\ -1 & (4 - \sqrt{7}) / 3 \\ 1 & (-4 + \sqrt{7}) / 3 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

(2) La norme  $l_\infty$  dans  $\mathbb{R}^n$  peut être définie par la matrice

$$B = \begin{pmatrix} I_n \\ -I_n \end{pmatrix} \quad \text{et} \quad b = (1, \dots, 1)^T$$

où  $I_n$  désigne la matrice unité de  $\mathbb{R}^n$ .

- (3) La norme  $l_1$  dans  $\mathbb{R}^n$  peut être définie par la matrice  $B(2^n \times n)$  dont les éléments de chaque ligne sont égaux à 1 en valeur absolue. Les  $2^n$  lignes sont obtenues en affectant ces valeurs absolues des signes + ou - et ce, de toutes les manières possibles.

Dans  $\mathbb{R}^3$ , nous aurions

$$B(8, 3) \equiv \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \quad \text{et} \quad b^T = (1, \dots, 1) \quad .$$

## § 2. APPROXIMATION DISCRETE

---

### 2.1. Définition.

---

Nous appellerons *problème d'approximation linéaire discret* tout problème de la forme

$$\min_{h \in \mathbb{R}^p} \|f + A h\|_B \quad (2.1)$$

où  $B(m, n)$  définit une norme polyédrale désignée par  $\|\cdot\|_B$  et où  $A(n, p)$  et  $f(n, 1)$  sont des données.

Nous supposerons également que

$$\text{rg}(A) = p$$

### 2.2. Caractérisation d'une Solution Optimale du Problème (2.1).

---

Remarque.

---

Nous utiliserons ici de nombreux résultats de la programmation linéaire. Afin de ne pas surcharger la présentation, nous renvoyons le lecteur à l'Appendice 1 où les théorèmes et les résultats employés seront cités et, si leur preuve ne réclame pas trop de présupposés, démontrés.

Réécrivons (2.1) sous la forme d'un programme linéaire classique. Le problème (2.1) est, en effet, trivialement équivalent à

$$\min_{\substack{h \in \mathbb{R}^p \\ \lambda \in \mathbb{R}}} [0, \dots, 0, 1] \begin{pmatrix} h \\ \lambda \end{pmatrix}$$

où  $(h, \lambda)$  sont soumis aux contraintes

$$\begin{pmatrix} B & A \\ (m, p+1) \end{pmatrix} \begin{pmatrix} h \\ \lambda \end{pmatrix} \leq -B f \quad (2.2)$$

Ecrivons également le dual de (2.2) [cf. Appendice 1.1]

$$\max_{\mu \in \mathbb{R}^m} [\mu_1, \dots, \mu_m] \begin{pmatrix} B & f \\ (m, 1) \end{pmatrix} \quad (2.3)$$

où  $\mu$  est soumis aux contraintes

$$\begin{aligned} & [\mu_1, \dots, \mu_m] \begin{pmatrix} -B & A \\ (1, m) \end{pmatrix} \begin{pmatrix} b \\ (m, p+1) \end{pmatrix} = [0, \dots, 0, 1] \begin{pmatrix} (1, p+1) \end{pmatrix} \\ & \mu_i \geq 0 \quad \forall i \in \{1, \dots, m\} \end{aligned}$$

Lemme 2.1.

Sous les hypothèses ci-dessus, la matrice  $\begin{pmatrix} B & A \\ (m, p+1) \end{pmatrix}$  est de rang  $p+1$ .

Preuve.

Puisque  $A(n, p)$  et  $\text{rg}(A) = p$  et que  $B(m, n)$  et  $\text{rg}(B) = n$ , on a

$$\text{rg}(B A) = p$$

Soient  $\rho_{i_1}(B A), \dots, \rho_{i_p}(B A)$   $p$  lignes de  $B A$  telles que la matrice

$$\begin{pmatrix} (B A)_{i_1 1} & \dots & (B A)_{i_1 p} \\ \vdots & & \vdots \\ (B A)_{i_p 1} & \dots & (B A)_{i_p p} \end{pmatrix}$$

soit de rang  $p$ .

La remarque (2), page 1.6, permet d'affirmer que la ligne  $-\rho_{i_p}(B A)$  appartient à  $B A$ . Dès lors, la matrice

$$\begin{pmatrix} (B A)_{i_1 1} & \dots & (B A)_{i_1 p} & -1 \\ \vdots & & \vdots & \vdots \\ (B A)_{i_p 1} & \dots & (B A)_{i_p p} & -1 \\ -(B A)_{i_p 1} & \dots & -(B A)_{i_p p} & -1 \end{pmatrix}$$

est immédiatement de rang  $p+1$ . Ceci prouve l'assertion puisque nous pouvons toujours supposer  $b_i = 1$ , pour tout  $i \in \{1, \dots, m\}$ .  $\square$

Pour continuer, introduisons quelques définitions et notations.

Soit  $\rho_i(\cdot)$  la  $i$ -ième ligne d'une matrice  $(\cdot)$ , transposée.

Nous dirons que le vecteur  $[h^*, \lambda^*] \in \mathbb{R}^{p+1}$  *sature* la  $i$ -ième contrainte de (2.2) si et seulement si  $\rho_i(B A) h^* - b_i \lambda^* = -(B f)_i$ .

Notons encore par  $\sigma_k$  un ensemble formé de  $p+1$  indices appartenant à  $\{1, \dots, m\}$ ,  $k$  étant un entier variant parmi les  $\binom{m}{p+1}$  sous-ensembles

possibles.

Cet ensemble d'indices jouant le même rôle que la référence classique dans l'algorithme d'échange pour le cas de la norme  $l_\infty$  (voir [5]), nous lui donnerons également le nom de *référence*.

Quelques résultats de programmation linéaire.

i)  $[h^*, \lambda^*]$  est un point optimal <sup>de base</sup> pour (2.2) ~~si et seulement si~~ :

(1)  $[h^*, \lambda^*]$  est admissible pour (2.2) ; (2.4)

(2) il existe une référence  $\sigma_k$  telle que les lignes correspondant aux indices appartenant à  $\sigma_k$  sont linéairement indépendantes ; (2.5)

(3)  $[h^*, \lambda^*]$  *sature* ces lignes, c'est-à-dire, pour tout  $j \in \sigma_k$ , (2.6)

$$\rho_j (BA) h - b_j \lambda = -(Bf)_j .$$

Désignons par  $B^{\sigma_k}$  la sous-matrice de  $B$  formée des  $p+1$  lignes composant  $\sigma_k$ . Nous avons donc

$$[-B^{\sigma_k} A, b^{\sigma_k}] \begin{pmatrix} h \\ \lambda \end{pmatrix} = B^{\sigma_k} f . \quad (2.7)$$

(Cf. Appendice 1.2 pour une présentation plus complète de ces résultats).

ii) D'autre part,  $[\mu_1^*, \dots, \mu_m^*]$  est une solution optimale pour le problème dual (2.3) si et seulement si (cf. Appendice 1.3) :

(1)  $[\mu_1^*, \dots, \mu_m^*]$  est admissible pour (2.3) ; (2.8)



- (2) les fonctions-objectif primale et duale sont égales à l'optimum, c'est-à-dire

$$\lambda^* = \sum_{i=1}^m \mu_i^* (B f)_i \quad (2.9)$$

iii) Il est aussi bien connu (cf. Appendice 1.2) qu'il existe une référence  $\sigma_j$  telle que :

$$(1) \quad [-B^{\sigma_j} A, b^{\sigma_j}] \text{ soit de rang } p+1 ; \quad (2.10)$$

$$(2) \quad [\mu_{i_1}^*, \dots, \mu_{i_{p+1}}^*] [-B^{\sigma_j} A, b^{\sigma_j}] = [0, \dots, 0, 1] \quad (2.11)$$

(p+1, p+1)

où les indices  $i_k$  sont les indices composant la référence  $\sigma_j$  ;

$$(3) \quad \text{les autres composantes de } [\mu_1^*, \dots, \mu_m^*] \text{ sont nulles} . \quad (2.12)$$

iv) Théorème des écarts complémentaires :

Si  $[h^*, \lambda^*]$  et  $[\mu_1^*, \dots, \mu_m^*]$  sont des solutions optimales pour les problèmes (2.2) et (2.3) respectivement, alors, pour tout  $i$  appartenant à  $\{1, \dots, m\}$ , on a

$$\mu_i^* [-\rho_i (B A) h^* + b_i - \rho_i (B) f] = 0 \quad (2.13)$$

Ces rappels de programmation linéaire vont nous permettre de démontrer un théorème qui est le théorème fondamental du développement de notre algorithme d'échange.

Théorème 2.1.

Le vecteur  $[h^*, \lambda^*]$  admissible pour le problème (2.2) est optimal pour ce problème si et seulement si il existe une référence  $\sigma_k$  et  $(p+1)$  multi-

plicateurs  $[\mu_{\lambda_1}^*, \dots, \mu_{\lambda_{p+1}}^*]$  tels que :

$$(1) \quad [-B^{\sigma_k} A, b^{\sigma_k}] \text{ soit de rang } p+1 \quad ; \quad (2.14)$$

$$(2) \quad [-B^{\sigma_k} A, b^{\sigma_k}] \begin{pmatrix} h^* \\ \lambda^* \end{pmatrix} = B^{\sigma_k} f \quad ; \quad (2.15)$$

$$(3) \quad \mu_{\lambda_j}^* \geq 0 \quad \forall j \in \{1, \dots, p+1\} \text{ et} \quad (2.16)$$

$$[0, \dots, 0, 1] = \sum_{j \in \sigma_k} \mu_j^* [-\rho_j(B A), b_j] ; \quad (2.17)$$

$$(4) \quad \lambda^* = \sum_{j \in \sigma_k} \mu_j^* \rho_j(B) f \quad . \quad (2.18)$$

Nécessité :

Utilisant (2.8) , (2.10) et (2.11) , nous pouvons construire une référence  $\sigma_j$  et un vecteur  $\mu^*$  dual optimal associé tels que

$$(1) \quad [-B^{\sigma_j} A, b^{\sigma_j}] \text{ soit de rang } p+1 \quad ; \quad (2.19)$$

$$(2) \quad \mu_l^* \geq 0 \quad \forall l \in \sigma_j \quad . \quad (2.20)$$

Par (2.12) , deux cas sont possibles :

(1) Il existe exactement  $p+1$  composantes  $\mu_j^*$  non nulles :

Par le théorème des écarts complémentaires, on a

$$\mu_l^* [-\rho_l(B) A h^* + b_l \lambda^* - \rho_l(B) f] = 0 \quad \forall l \in \{1, \dots, m\}$$

comme  $\mu_l^* > 0$  pour tout  $l \in \sigma_j$  .



Nous pouvons donc prendre  $\sigma_k = \sigma_j$ .

La relation (2.17) découle de (2.11) et la relation (2.18) de (2.9).

(2) Il existe moins de  $(p+1)$  composantes  $\mu_j^*$  non nulles :

D'après (2.5), (2.6) et (2.7), il existe une référence  $\sigma_1$  telle que

$$(1) \quad [-B^{\sigma_1} A, b^{\sigma_1}] \text{ soit de rang } p+1 \quad ;$$

$$(2) \quad [-B^{\sigma_1} A, b^{\sigma_1}] \begin{pmatrix} h^* \\ \lambda^* \end{pmatrix} = -B^{\sigma_1} f \quad .$$

Construisons la référence  $\sigma_k$  de la manière suivante :

$$\mu_i^* > 0 \Rightarrow i \in \sigma_k \quad .$$

Nous obtenons ainsi une référence formée de moins de  $(p+1)$  indices dont les lignes sont linéairement indépendantes et, par le Théorème des Ecartés Complémentaires, sont saturées par  $[h^*, \lambda^*]$ .

Il nous suffit, dès lors, de compléter cette référence par des lignes de  $\sigma_1$  telles que la matrice

$$[-B^{\sigma_k} A, b^{\sigma_k}] \text{ soit de rang } p+1 \quad ,$$

ceci étant toujours possible, et de poser les  $\mu_i^*$  correspondants égaux à 0.

Nous avons, dès lors, obtenu une référence  $\sigma_k$  vérifiant les conditions (2.14) à (2.18).

Suffisance :

Considérons le vecteur  $(\mu_1^*, \dots, \mu_n^*)$  tel que

$$i \notin \sigma_k \Rightarrow \mu_i^* = 0$$

$$i \in \sigma_k \Rightarrow \mu_i^* = \mu_i^*$$

Nous avons ainsi construit une solution  $[\mu_1^*, \dots, \mu_n^*]$ , duale admissible grâce à (2.17) .

D'autre part, d'après (2.18) ,

$$\lambda^* = \sum_{j \in \sigma_k} \mu_j^* \rho_j(B) f$$

Ceci, compte tenu de (2.8) et (2.9) , assure l'optimalité de  $[h^*, \lambda^*]$  .  $\square$

### § 3. GENERALISATION AUX PROBLEMES AVEC CONTRAINTES

---

#### 3.1. Définition.

---

Nous allons généraliser le problème (2.1) au cas où le vecteur solution  $h^*$  est soumis à des contraintes linéaires

$$C h \leq d$$

Pour ce faire, considérons une matrice  $A(n, p)$ , de rang  $p$ , donnée ,  
 une matrice  $C(q, p)$ , donnée ,  
 un vecteur  $d$ , donné ,  
 un vecteur  $f$ , donné ,

et envisageons le problème :

$$\left. \begin{array}{l} \text{minimiser } \|r(h)\|_B \\ h \in \mathbb{R}^p \\ \text{où } r = A h + f \\ \text{mais soumis aux contraintes} \\ C h \leq d \\ \text{les contraintes étant supposées consistantes.} \end{array} \right\} \quad (3.1)$$

Nous pouvons réécrire ce problème sous la forme du programme linéaire :

$$\left. \begin{array}{l} \min [0, \dots, 0, 1] \begin{pmatrix} h \\ \lambda \end{pmatrix} \\ \text{sous les contraintes} \end{array} \right\} \quad (3.2)$$



$\mu_{i_s}^*$  ( $s = \# \sigma_{1i}$ ) et  $[\eta_{i_1}^*, \dots, \eta_{i_{p+1-s}}^*]$  tels que

$$(1) \quad \begin{pmatrix} -B^{\sigma_{1i}} A, b^{\sigma_{1i}} \\ -C^{\sigma_{2i}}, 0 \end{pmatrix} \text{ soit de rang } p+1 ;$$

$$(2) \quad \begin{pmatrix} -B^{\sigma_{1i}} A, b^{\sigma_{1i}} \\ -C^{\sigma_{2i}}, 0 \end{pmatrix} \begin{pmatrix} h^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} B^{\sigma_{1i}} b \\ -d^{\sigma_{2i}} \end{pmatrix} ; \quad (3.4)$$

$$(3) \quad \mu_{ij}^* \geq 0 \quad \forall j \in \{1, \dots, s\} \text{ et}$$

$$\eta_{ij}^* \geq 0 \quad \forall j \in \{1, \dots, p+1-s\} \quad (3.5)$$

$$\text{et} \quad (3.6)$$

$$\begin{aligned} [0, \dots, 0, 1] = & \sum_{j \in \sigma_{1i}} \mu_j^* [-\rho_j(BA), b_j] \\ & + \sum_{j \in \sigma_{2i}} \eta_j^* [-\rho_j(C), 0] ; \end{aligned}$$

$$(4) \quad \lambda^* = \sum_{j \in \sigma_{1i}} \mu_j^* \rho_j(B) b - \sum_{j \in \sigma_{2i}} \eta_j^* d_j . \quad (3.7)$$

## § 4. UN ALGORITHME D'ECHANGE

---

### 4.1. Notion Intuitive.

---

L'algorithme que nous allons développer a pour but de résoudre le problème (3.1).

Afin de bien comprendre sa signification, réécrivons le problème (3.1) et son dual.

(3.1) s'écrit

$$\left. \begin{array}{l}
 \min [0, \dots, 0, 1] \begin{pmatrix} h \\ \lambda \end{pmatrix} \\
 \text{soumis aux contraintes} \\
 \begin{pmatrix} -B & A & , & b \\ -C & , & 0 \end{pmatrix} \begin{pmatrix} h \\ \lambda \end{pmatrix} \geq \begin{pmatrix} B & f \\ -d \end{pmatrix} \\
 (q+m, p+1)
 \end{array} \right\} \quad (4.1)$$

et son dual

$$\left. \begin{array}{l}
 \max [v_1, \dots, v_{q+m}] \begin{pmatrix} B & f \\ -d \end{pmatrix} \\
 (q+m, 1) \\
 \text{soumis aux contraintes}
 \end{array} \right\}$$

$$\left. \begin{aligned} [v_1, \dots, v_{q+m}] \begin{pmatrix} -B A & b \\ -C & 0 \end{pmatrix} &= [0, \dots, 0, 1] \\ v_i &\geq 0 \quad \forall i = 1, \dots, m+q \end{aligned} \right\} \quad (4.2)$$

Nous allons nous déplacer de solution duale admissible en solution duale admissible de sorte que, à chaque itération, l'on fait correspondre à la solution duale une solution primale non-admissible et un coût de la fonction-objectif primale égale au coût de la fonction-objectif duale. On accroît alors ce coût jusqu'au moment où la solution primale devient admissible, c'est-à-dire où l'on est à l'optimum.

Afin de ne pas alourdir la description de l'algorithme par des démonstrations de points très techniques, ceux-ci seront explicités après.

Une étoile suivie d'un chiffre renvoie donc à l'explication de ce point après la description de l'algorithme.

#### 4.2. Algorithme - Description.

---

Etape 0 : Initialisation : \* - 1 .

---

Le départ de l'algorithme réclame la construction d'une référence initiale ainsi que des multiplicateurs associés de telle manière que les relations (3.5)-(3.6) soient satisfaites.



Nous montrerons ci-après la manière de déterminer cette référence initiale. Mentionnons cependant que, dans la suite, l'algorithme permettra d'assurer que toutes les références utilisées vérifient ces relations.

Nous prendrons, pour débiter,  $\sigma_{21} = \emptyset$  et un ensemble  $\sigma_{11}$  formé de  $p+1$  indices tel que

$$\left. \begin{aligned}
 & i := 1 \\
 & \left. \begin{aligned}
 & B^{\sigma_{11}} A \text{ soit de rang } p \\
 & (p+1, p) \\
 & \text{et} \\
 & \sum_{j \in \sigma_{11}} \xi_j^{(i)} \rho_j(B) A = (0, \dots, 0) \\
 & \text{et} \\
 & \xi_j^{(i)} \geq 0 \quad \forall j \in \sigma_{11} \\
 & \text{et} \\
 & \sum_{j \in \sigma_{11}} b_j \xi_j^{(i)} = 1
 \end{aligned} \right\} \quad (4.3)
 \end{aligned} \right\}$$

Etape 1.

A l'étape  $i$ , nous supposons que nous disposons de deux ensembles d'indices  $\sigma_{1i}$  et  $\sigma_{2i}$  vérifiant les conditions suivantes :

$$\text{La matrice } Z_i = \begin{pmatrix} B^{\sigma_{1i}} A \\ C^{\sigma_{2i}} \end{pmatrix}^* \text{ est de rang } p ;$$



l'étoile indique que les lignes de  $Z_i$  sont celles de  $\begin{pmatrix} B^{\sigma_{1i}} & A \\ C^{\sigma_{2i}} \end{pmatrix}$ , mais dans un ordre qui peut être différent

$$\left. \begin{aligned} &\text{Il existe des nombres } \xi_j^{(i)}, j \in \sigma_{1i} \text{ et } \eta_j^{(i)}, j \in \sigma_{2i}, \\ &\text{positifs ou nuls tels que} \\ &\sum_{j \in \sigma_{1i}} b_j \xi_j^{(i)} = 1 \\ &\sum_{j \in \sigma_{1i}} \xi_j^{(i)} \rho_j(B) A + \sum_{j \in \sigma_{2i}} \eta_j^{(i)} \rho_j(C) = (0, \dots, 0) . \end{aligned} \right\} (4.4)$$

Nous rassemblons les multiplicateurs  $\xi_j^{(i)}$  et  $\eta_j^{(i)}$  dans un vecteur  $\omega_i$  tel que

$$\omega_i^T Z_i = (0, \dots, 0) .$$

Remarquons que l'étape d'initialisation permet de satisfaire toutes ces conditions durant la première itération ( $\sigma_{2i} = \emptyset$ ). Il est également clair que nous avons ainsi construit une référence vérifiant les relations (3.5) et (3.6).

#### Etape 2.

Cette étape va déterminer une approximation de la norme par la relation (3.7) et chercher un vecteur  $h^{(i)}$  associé à cette référence et à cette norme grâce à la relation (3.4). Pour ce faire, factorisons  $Z_i$  sous la forme

$$\begin{matrix} Z_i \\ (p+1, p) \end{matrix} = \begin{matrix} Q_i \\ (p+1, p+1) \end{matrix} \begin{matrix} \begin{pmatrix} U_i \\ 0 \end{pmatrix} \\ (p+1, p) \end{matrix}$$

où  $Q_i$  est orthogonale et  $U_i(p, p)$  est triangulaire supérieure.

Ceci est toujours possible (voir \* - 2).

Calculons ensuite la valeur de la norme associée à la référence utilisée si celle-ci était optimale. La formule (3.7) donne

$$\lambda^{(i)} = \sum_{j \in \sigma_{1i}} \xi_j^{(i)} \rho_j(B) f - \sum_{j \in \sigma_{2i}} \eta_j^{(i)} d_j \quad (4.5)$$

Nous déterminons alors le vecteur  $h^{(i)}$ , éventuelle solution optimale, par la relation (3.4) qui, par notre factorisation, est identique à

$$\begin{pmatrix} U_i \\ 0 \end{pmatrix} h^{(i)} = Q_i^T \begin{pmatrix} \lambda^{(i)} b^{\sigma_{1i}} - B^{\sigma_{2i}} f \\ d^{\sigma_{2i}} \end{pmatrix}^* \quad (* - 3)$$

et successivement

$$r^{(i)} = A h^{(i)} + f$$

$$s^{(i)} = C h^{(i)} - d$$

$$R^{(i)} = \max_{1 \leq j \leq m} \frac{\rho_j(B) r^{(i)}}{b_j} - \lambda^{(i)}$$

$$S^{(i)} = \max_{1 \leq j \leq q} s_j^{(i)}$$

### Etape 3.

Cette étape détermine si le vecteur  $[h^{(i)}, \lambda^{(i)}]$  est optimal. Ceci, étant donné les constructions ci-dessus, revient à vérifier si le vecteur  $[h^{(i)}, \lambda^{(i)}]$  est admissible pour (3.1). Dans le cas de non-admissibilité, nous

déterminons la ligne entrant dans la référence suivante.

Donc, si  $\max \{R^{(i)}, S^{(i)}\} \leq 0$ , *STOP*, on est à l'optimum.

Autrement, si

- (1)  $R^{(i)} > S^{(i)}$ , la ligne qui entre dans la nouvelle référence est  $v_i^T = \rho_1(B)$  où 1 est l'indice maximisant  $R^{(i)}$ .
- (2)  $R^{(i)} \leq S^{(i)}$ , la ligne qui entre dans la nouvelle référence est  $v_i^T = \rho_1(C)$  où 1 est l'indice maximisant  $S^{(i)}$ .

Etape 4.

Cette étape détermine la ligne quittant la référence; le schéma ci-dessous est surtout technique et vise à assurer, comme nous le verrons, la création d'une référence vérifiant les conditions de l'Etape 1.

Pour ce faire, considérons la matrice augmentée

$$\begin{pmatrix} z_i \\ v_i^T \end{pmatrix}_{(p+2, p)} = \begin{pmatrix} Q_i & 0 \\ (p+1, p+1) & (p+1, 1) \\ 0 & 1 \end{pmatrix}_{(1, p+1)} \begin{pmatrix} U_i \\ 0 \\ v_i^T \end{pmatrix}.$$

D'où

$$\begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = \begin{pmatrix} Q_i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_p \\ 0 \\ v_i^T U_i^{-1} \end{pmatrix} U_i. \quad (4.6)$$

Nous factorisons la matrice  $\begin{pmatrix} I_p \\ 0 \\ v_i^T U_i^{-1} \end{pmatrix}$  sous la forme

$$\begin{pmatrix} I_p \\ 0 \\ v_i^T U_i^{-1} \end{pmatrix} = \begin{pmatrix} p_1^T \\ \vdots \\ p_p^T \end{pmatrix} \begin{pmatrix} V \\ 0 \\ 0 \end{pmatrix}$$

où les  $p_i$  ( $p+2, p+2$ ) sont orthogonales et où  $V$  ( $p, p$ ) est triangulaire supérieure.

Ceci est toujours possible (voir \* - 4) .

Calculons

$$\begin{pmatrix} x_i \\ (p+2, 1) \end{pmatrix} = \begin{pmatrix} Q_i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_1^T \\ \vdots \\ p_p^T \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (*) - 5$$

avec

$$(x_i)_{p+2} > 0 \quad . \quad (*) - 6$$

Deux cas sont alors possibles :

- (1) Il existe  $j_0$  tel que  $(\omega_i)_{j_0} = 0$  et  $(x_i)_{j_0} < 0$  ;  
La ligne sortant de  $Z_i$  est alors  $\rho_{j_0}(z_i)$  .

- (2) Autrement, nous considérons

$$\gamma_i = \min_{(\omega_i)_j > 0} \frac{(x_i)_j}{(\omega_i)_j}$$

et soit  $j_0$  l'indice donnant le minimum;

la ligne sortant de  $Z_i$  est alors  $\rho_{j_0}(z_i)$ .

Etape 5.

Cette étape détermine la nouvelle matrice  $Z_{i+1}$  et le vecteur des multiplicateurs associés de telle manière que les hypothèses de l'Etape 1 soient satisfaites. A nouveau, ces considérations sont assez techniques et seront expliquées par après.

Posons

$$y_i = (w_i)_{j_0} x_i - (x_i)_{j_0} \begin{pmatrix} w_i \\ 0 \end{pmatrix} .$$

Nous avons

$$y_i \geq 0 \quad (\ast - 8)$$

et

$$(y_i)_{j_0} = 0 .$$

Posons

$$(\hat{y}_i)_j = (y_i)_j \quad j = 1, \dots, p+1 ; j \neq j_0$$

$$(\hat{y}_i)_{j_0} = (y_i)_{p+2} .$$

Construisons la nouvelle référence. A cette fin, posons

$$\sigma'_{li} = \begin{cases} \sigma_{li} \cup \{1\} & \text{si l'indice } l \text{ entrant dans la référence } \in B , \\ \sigma_{li} & \text{sinon} \end{cases} ;$$

$$\sigma'_{2i} = \begin{cases} \sigma_{2i} \cup \{1\} & \text{si l'indice } 1 \text{ entrant dans la référence } \in C, \\ \sigma_{2i} & \text{sinon} \end{cases}$$

Enfin,

$$\sigma_{1(i+1)} = \begin{cases} \sigma'_{1i} \setminus \{j_1\} & \text{si la ligne sortant de la référence est} \\ & \text{la ligne } \rho_{j_1}(B), \quad j_1 \in \{1, \dots, m\}, \\ \sigma'_{1i} & \text{sinon} \end{cases};$$

$$\sigma_{2(i+1)} = \begin{cases} \sigma'_{2i} \setminus \{j_1\} & \text{si la ligne sortant de la référence est} \\ & \text{la ligne } \rho_{j_1}(C), \quad j_1 \in \{1, \dots, m\}, \\ \sigma'_{2i} & \text{sinon} \end{cases}.$$

$$z_{i+1} = z_i + c_{j_0} [v_i^T - \rho_{j_0}(z_i)]$$

(\*-9)

où

$$c_k^T = (\delta_{ik})_{i=1, \dots, p+1}$$

$$w_{i+1} = \frac{\hat{y}_i}{\sum_{j \in \sigma_{1(i+1)}} b_j (\hat{y}_i)_j}$$

où le facteur ou dénominateur est un facteur de normalisation.

L'itération  $i$  est terminée.

Il suffit alors de poser  $i = i+1$  et de retourner à l'Etape 1.



Discussion de certains points de l'algorithme.

---

(\*-1).

---

L'initialisation de l'algorithme réclame la construction d'une matrice  $B^{\sigma_{1i}} A$  de rang  $p$  telle que

$$(0 \quad \dots \quad 0) = \sum_{j \in \sigma_{10}} \xi_j^{(0)} \rho_j(B) A$$

et

$$\xi_j^{(0)} \geq 0 \quad \forall j \in \sigma_{10}.$$

Ceci est toujours possible si  $A$  est de rang  $p$ . Pour ce faire, considérons une matrice  $B_1(n, n)$  de rang  $n$ .

Nous pré-multiplions  $A$  par  $B_1$  et pivotons sur les lignes ou colonnes de  $B_1 A$  afin de trouver  $p$  lignes de  $B_1 A$  linéairement indépendantes.

Soient  $\rho_{i1}(B_1 A)$ ,  $\dots$ ,  $\rho_{ip}(B_1 A)$  ces lignes.

La matrice  $B^{\sigma_{10}}$  est alors égale à

$$\begin{pmatrix} \rho_{i1}(B) \\ \vdots \\ \rho_{ip}(B) \\ s_i^T \end{pmatrix}$$

où  $s_i^T$  est une ligne quelconque de  $B$ .

Une élimination de Gauss permet alors de déterminer  $(\xi_1^{(0)}, \dots, \xi_{(p)}^0)$  tels que



$$[\xi_1^{(0)}, \dots, \xi_p^{(0)}, 1] B^{\sigma_{1i}} A = [0 \quad \dots \quad 0]$$

Si, par hasard, il existe  $k$  tel que  $\xi_k < 0$ , il suffit alors de poser

$$\rho_{i_k}(B) = -\rho_{i_k}(B)$$

et

$$\xi_k^{(0)} = -\xi_k^0$$

Une normalisation des  $\xi_j^{(0)}$  permet de satisfaire les conditions (4.3).

(\*-2).

Nous ne démontrerons pas la possibilité ni la construction de cette factorisation; pour plus de détails, voir [4, pp. 59-60].

(\*-3).

Le système linéaire (\*-3) déterminant  $h^{(i)}$  est un système de  $(p+1)$  équations à  $p$  inconnues ( $\dim h^{(i)} = p$ ); vérifions qu'il est compatible.

Le système est équivalent à

$$Z_i h^{(i)} = \begin{pmatrix} \lambda^{(i)} b^{\sigma_{1i}} - B^{\sigma_{1i}} f \\ d^{\sigma_{2i}} \end{pmatrix}^*$$

Supposons

$$(w_i)_k \neq 0$$

et

7

$$\left( \begin{array}{ccc} \lambda^{(i)} & b^{\sigma_{1i}} & -B^{\sigma_{1i}} f \\ & d^{\sigma_{2i}} & \end{array} \right)_k^* = d_1$$

Par les relations (4.4), nous avons

$$\rho_k(z_i) = -\frac{1}{(w_i)_k} \left( \sum_{j \in \sigma_{1i}} \rho_j(z_i) \xi_j^{(i)} + \sum_{\substack{j \in \sigma_{2i} \\ j \neq 1}} \rho_j(z_i) \eta_j^{(i)} \right)$$

et, d'après la relation (4.5),

$$d_1 = -\frac{1}{(w_i)_k} \left( \sum_{j \in \sigma_{1i}} (\lambda^{(i)} - \xi_j^{(i)} \rho_j(B) f) + \sum_{\substack{j \in \sigma_{2i} \\ j \neq 1}} \eta_j^{(i)} d_j \right),$$

ce qui nous prouve la compatibilité du système (\*-3).

(\*-4).

Montrons que la factorisation

$$\begin{pmatrix} I \\ 0 \\ V_i^T U_i^{-1} \end{pmatrix} = P_1^T \dots P_p^T \begin{pmatrix} V \\ 0 \\ 0 \end{pmatrix}$$

est possible.

Pour ce faire, posons

$$P_k = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & \cos \theta & & & -\sin \theta & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ & & \sin \theta & & & & & \cos \theta \end{pmatrix} \leftarrow k$$

où les éléments hors diagonales non-marqués sont nuls.

Démontrons l'assertion par récurrence et, pour cela, supposons qu'à l'étape  $k$ ,

la matrice  $\begin{pmatrix} I \\ 0 \\ V_i^T U_i^{-1} \end{pmatrix}$  ait été amenée sous la forme :

$$\begin{pmatrix} v_{11}^{(k)} & v_{12}^{(k)} & \dots & \dots & \dots & \dots & \dots & \dots & v_{1,p}^{(k)} \\ & v_{22}^{(k)} & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ & & \ddots & \dots & \dots & \dots & \dots & \dots & v_{k-1,p}^{(k)} \\ & & & v_{k-1,k-1}^{(k)} & \dots & \dots & \dots & \dots & 0 \\ & & & & 1 & 0 & \dots & \dots & \vdots \\ & & & & & 1 & \dots & \dots & \vdots \\ & & & & & & \ddots & 0 & 1 \\ & & & & & & & 1 & \vdots \\ & & & & & & & & 1 \\ 0 & \dots & \dots & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & v_{p+2,k}^{(k)} & \dots & \dots & \dots & v_{p+2,p}^{(k)} \end{pmatrix} = V_k \cdot$$

Appliquant  $P_k$  à cette matrice, nous obtenons

$$\sin \theta + \cos \theta v_{p+2,k}^{(k)} = v_{p+2,k}^{(k+1)}.$$

Dès lors,  $v_{p+2,k}^{(k+1)} = 0$  si et seulement si

$$\operatorname{tg} \theta = -v_{p+2,k}^{(k)}.$$

Vérifions que l'application de  $P_k$  à la matrice  $V_k$  n'en modifie pas la partie hachurée et les zéros des dernière et avant-dernière lignes. On a, en effet,

$$v_{ij}^{(k+1)} = \sum_{l=1}^{p+2} p_{il}^{(k)} v_{lj}^{(k)}.$$

(1)  $i = 1, \dots, p$  et  $j < i$  (partie hachurée) :

$$p_{i1}^{(k)} = 0$$

sauf si  $i = 1$ , mais alors  $v_{ij}^{(k)} = v_{1j}^{(k)} = 0$  car  $j < i = 1$ ,  
 si  $i = k$  et  $l = p+2$ , or  $i = k \Rightarrow j < i \Rightarrow v_{p+2,j}^{(k)} = 0$ .

(2)  $i = p+1$  (avant dernière ligne) :

Seule la composante  $p_{p+1,p+1}^{(k)}$  est différente de zéro, mais  
 $v_{p+1,j} = 0$  pour tout  $j$ .

(3)  $i = p+2$  et  $j < k$  (dernière ligne) :

$$p_{p+2,1}^{(k)} = 0$$

sauf si  $l = k \Rightarrow v_{1j}^{(k)} = v_{kj}^{(k)} = 0$  car  $j < k$ ,  
 si  $l = p+2 \Rightarrow v_{p+2,j}^{(k)} = 0$  car  $j < k$ .

(\*-6).

---

Vérifions que l'on peut toujours s'arranger pour que  $(x_i)_{p+2} > 0$ .

En effet, la définition de  $x_i$  et un calcul élémentaire montrent que

$$x_i = \begin{pmatrix} \sin \theta_2 \prod_{i=2}^p \cos \theta_i \\ \sin \theta_2 \prod_{i=3}^p \cos \theta_i \\ \vdots \\ \sin \theta_p \\ 0 \\ \prod_{i=1}^p \cos \theta_i \end{pmatrix}$$

Ceci permet de calculer  $x_i$  récursivement sans stocker les  $P_i$  et d'assurer, en choisissant  $\theta_i$  ou  $\theta_i + \pi$ ,  $\cos \theta_i > 0$  pour tout  $i$ .

Remarquons aussi que, puisque

$$\operatorname{tg} \theta_j = -v_{p+2,j}^{(j)},$$

on a toujours  $\cos \theta_j \neq 0$ .

(\*-8).

---

Vérifions que le vecteur  $y_i$  construit à l'étape 5 est non-négatif. Deux cas sont à envisager :

(1) Il existe un indice  $j_0$  tel que  $w_{j_0} = 0$  et  $x_{j_0} < 0$  : c'est trivial;

(2) Autrement, il nous faut voir que, pour tout  $j \in \{1, \dots, p+1\}$ ,

$$(y_i)_j \geq 0$$

où

$$(y_i)_j = w_{j_0} (x_i)_j - x_{j_0} (w_i)_j \geq 0$$

$$\Leftrightarrow w_{j_0} (x_i)_j \geq x_{j_0} (w_i)_j \quad (4.7)$$

$$i) \quad (w_i)_j = 0 \Rightarrow (x_i)_j \geq 0 \Rightarrow (y_i)_j \geq 0 \quad ;$$

ii) autrement, (4.7) est vérifiée si et seulement si

$$\frac{(x_i)_j}{(w_i)_j} \geq \frac{(x_i)_{j_0}}{(w_i)_{j_0}} \quad ,$$

ce qui est une conséquence du choix de l'indice  $j_0$ .

(\*-9).

Il nous faut vérifier que nous avons construit deux nouveaux ensembles d'indices  $\sigma_{1(i+1)}$  et  $\sigma_{2(i+1)}$  vérifiant les hypothèses de l'Etape 1, essentiellement :

$$\text{rg}(z_{i+1}) = p$$

$$w_{i+1} z_{i+1} = (0, \dots, 0)$$

$$w_{i+1} \geq 0$$

$$\sum_{j \in \sigma_{1i}} b_j \xi_j^{(i)} = 1$$

Les deux dernières relations sont des conséquences directes de  $(*-6)$ ,  $(*-8)$

et de la construction de  $w_{i+1}$ .

Pour démontrer que  $\text{rg}(z_{i+1}) = p$ , nous avons besoin de deux résultats :

$$(1) \quad x_i^T \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = [0, \dots, 0] : \quad (4.8)$$

Cette relation d'orthogonalité va d'ailleurs justifier la construction assez arbitraire du vecteur  $x_i$ .

Nous avons, en effet,

$$x_i^T \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = [0, \dots, 0, 1] \cdot P_p \dots P_1 \begin{pmatrix} Q_i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_i \\ v_i^T \end{pmatrix}.$$

$(1, p+2) \quad (p+2, p) \quad (1, p+2) \quad (p+2, p+2) \quad (p+2, p+2) \quad (p+2, p)$

Utilisant (4.8) et (\*-4), ceci équivaut à

$$x_i^T \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = [0, \dots, 0, 1] \begin{pmatrix} V \\ 0 \\ 0 \end{pmatrix} U_i$$

$(1, p+2) \quad (p+2, p) \quad (p, p)$

ou encore

$$x_i^T \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = [0, \dots, 0]$$

puisque  $V(p, p)$  et  $U_i(p, p)$  sont triangulaires supérieurs.

(2) Lemme.:

Si  $a_1, \dots, a_{p+1}$  forment une partie génératrice de  $\mathbb{R}^p$  et si



$$a = \sum_{i=1}^{p+1} \lambda_i a_i$$

avec  $\lambda_k \neq 0$  et  $k \in \{1, \dots, p+1\}$ , alors  $a_1, \dots, a_{k-1}, a, a_{k+1}, \dots, a_{p+1}$  forment une partie génératrice de  $\mathbb{R}^p$ .

Preuve :

Il nous suffit de montrer que, pour tout  $x \in \mathbb{R}^p$ , il existe  $\mu_1, \dots, \mu_{p+1}$  tels que

$$x = \sum_{\substack{i=1 \\ i \neq k}}^{p+1} \mu_i a_i + \mu_k a$$

Or, pour tout  $x \in \mathbb{R}^p$ , il existe  $v_1, \dots, v_{p+1}$  tels que

$$x = \sum_{i=1}^{p+1} v_i a_i$$

De plus,

$$a_k = \frac{1}{\lambda_k} \left[ a - \sum_{\substack{i=1 \\ i \neq k}}^{p+1} \lambda_i a_i \right],$$

ce qui entraîne

$$x = \sum_{i=1}^{p+1} \left[ v_i - \frac{\lambda_i}{\lambda_k} v_k \right] a_i + \frac{v_k}{\lambda_k} a \quad . \square$$

Revenons à notre thèse, à savoir  $\text{rg}(Z_{i+1}) = p$ .

Nous avons obtenu deux relations d'orthogonalité :

$$x_i^T \begin{pmatrix} Z_i \\ v_i^T \end{pmatrix} = [0 \quad \dots \quad 0]$$

et

$$[w_i^T, 0] \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} = [0 \quad \dots \quad 0] \quad . \quad (4.9)$$

De ces deux relations de dépendance linéaire, nous tirons

$$v_i^T = - \frac{1}{(x_i)_{p+2}} \sum_{j=1}^{p+1} \rho_j(z_i) (x_i)_j$$

et

$$v_i^T = \sum_{j=1}^{p+1} [(w_i)_j - \frac{(x_i)_j}{(x_i)_{p+2}}] \rho_j(z_i) \quad . \quad (4.10)$$

Dès lors, ou bien  $(x_i)_{j_0} \neq 0$  et, par le Lemme,  $z_{i+1}$  est de rang  $p$ ,  
ou bien  $(x_i)_{j_0} = 0$  et, par le choix de  $j_0$ , on a  $(w_i)_{j_0} > 0$ .

La relation (4.10) et le lemme impliquent à nouveau que  $\text{rg}(z_{i+1}) = p$ .

Pour terminer la discussion des points de l'algorithme, il nous faut montrer que  $w_{i+1}^T z_{i+1} = 0$ .

Or,

$$\begin{aligned} w_{i+1}^T z_{i+1} &= \sum_{j=1}^{p+1} (w_{i+1})_j \rho_j(z_{i+1}) \\ &= K \sum_{\substack{j=1 \\ j \neq j_0}}^{p+1} [(x_i)_j (w_i)_{j_0} - (w_i)_j (x_i)_{j_0}] \rho_j(z_i) + K (y_i)_{p+2} v_i^T \end{aligned}$$

où  $K$  est la constante de normalisation égale à

$$1 / \sum_{j \in \sigma_1(i+1)} b_j (\hat{y}_i)_j$$

$$\begin{aligned}
&= K \sum_{\substack{j=1 \\ j \neq j_0}}^{p+1} [(x_i)_j (w_i)_{j_0} - (w_i)_j (x_i)_{j_0}] \rho_j(z_i) + K (y_i)_{p+2} v_i^T \\
&\quad + (x_i)_{j_0} (w_i)_{j_0} (\rho_{j_0}(z_i) - \rho_{j_0}(z_i)) K \\
&= K \left[ (w_i)_{j_0} x_i^T \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} - (x_i)_{j_0} (w_i^T, 0) \begin{pmatrix} z_i \\ v_i^T \end{pmatrix} \right] \\
&= 0
\end{aligned}$$

Ce dernier résultat est obtenu par (4.8) et par hypothèse de récurrence.

## § 5. CONVERGENCE DE L'ALGORITHME

---

Dans l'algorithme que nous avons construit, nous nous déplaçons de référence en référence de telle manière que, à chaque itération, les conditions (3.4)-(3.7) soient vérifiées.

Nous allons montrer que la valeur  $\lambda^*$  optimale est supérieure à toute approximation  $\lambda^{(i)}$  construite par l'algorithme et que, d'autre part,  $\lambda^{(i+1)} > \lambda^{(i)}$  prouvera qu'une même référence ne peut apparaître plus d'une fois durant l'algorithme.

Le nombre de références étant fini, nous obtiendrons finalement une référence telle que le vecteur associé  $(h^{(s)}, \lambda^{(s)})$  soit admissible et donc optimal.

- (1) Prouvons que toute approximation  $(h^{(i)}, \lambda^{(i)})$  construite par l'algorithme est telle que

$$\lambda^* \geq \lambda^{(s)} \quad ;$$

l'égalité a lieu à l'optimum.

En effet, le vecteur  $w_i$ , complété par les zéros adéquats, forme une solution admissible du problème dual dont le coût est égal à  $\lambda^{(i)}$ .

Or, un théorème bien connu de programmation linéaire [cf. Appendice 1.3] assure que la valeur de la fonction objectif duale est inférieure ou égale à la valeur de la fonction objectif primale et lui est égale à l'optimum.  $\square$

(2) Si  $w_{i+1} > 0$ , alors  $\lambda^{(i+1)} > \lambda^{(i)}$ .

Par (4.5), nous avons

$$\lambda^{(i+1)} = \sum_{j \in \sigma_1(i+1)} \xi_j^{(i+1)} \rho_j(B) f - \sum_{j \in \sigma_2(i+1)} \eta_j^{(i+1)} d_j. \quad (5.1)$$

Par (4.4), nous obtenons

$$\begin{aligned} [0 \quad \dots \quad 0] &= \sum_{j \in \sigma_1(i+1)} \xi_j^{(i+1)} \rho_j(B) A \\ &\quad + \sum_{j \in \sigma_2(i+1)} \eta_j^{(i+1)} \rho_j(C). \end{aligned}$$

Dès lors, (5.1) est équivalent à

$$\begin{aligned} \lambda^{(i+1)} &= \sum_{j \in \sigma_1(i+1)} \xi_j^{(i+1)} \rho_j(B) (f + A h^{(i)}) \\ &\quad + \sum_{j \in \sigma_2(i+1)} \eta_j^{(i+1)} (\rho_j(C) h^{(i)} - d_j) \\ &= \lambda^{(i)} \sum_{j \in \sigma_1(i+1)} \sum_{\sigma_{1i}} \xi_j^{(i+1)} b_j \\ &\quad + \begin{cases} (\lambda^{(i)} + R^{(i)} b_1 \xi_1^{(i+1)}) & \text{si } 1 \in \sigma_1(i+1) \\ S^{(i)} \eta_1^{(i+1)} & \text{si } 1 \in \sigma_2(i+1) \end{cases} \end{aligned}$$

où 1 est l'indice de l'Etape 3 de l'algorithme

$$= \begin{cases} \lambda^{(i)} + R^{(i)} b_1 \xi_1^{(i+1)} & \text{si } 1 \in \sigma_1(i+1) \\ \lambda^{(i)} + S^{(i)} \eta_1^{(i+1)} & \text{si } 1 \in \sigma_2(i+1) \end{cases}$$

$$> \lambda^{(i)}$$

car

$$w_{i+1} > 0$$

implique

$$\xi_1^{(i+1)} > 0$$

Commentaires.:

La convergence de l'algorithme peut ne pas avoir lieu si  $\xi_1^{(i+1)} = 0$ . En effet, dans ce cas, il est possible de retrouver une référence déjà utilisée, ce qui entraînerait un cyclage de l'algorithme.

Notons cependant que cette possibilité de cyclage théorique semble être très peu probable sur le plan numérique. A titre de renseignement, signalons que trois cents utilisations de l'algorithme n'ont conduit à aucun cyclage.

Il est probable que des précautions inspirées du simplexe (lexicographiques, par exemple) permettent de résoudre ce problème. Cependant, comme il ne se pose pratiquement pas, nous n'avons pas développé cette amélioration possible.



## CONCLUSIONS

---

- (1) Il est très important de noter que notre algorithme généralise en fait l'algorithme d'échange de Stiefel dans le cas de la norme  $l_\infty$ .

Une description de celui-ci se trouve, par exemple, dans [6].

Compte-tenu de la description de notre algorithme et des deux relations d'orthogonalité ( (4.9) et (4.10) ), il n'y a aucune difficulté à montrer que notre algorithme suit, dans le cas de la norme  $l_\infty$ , les étapes de celui de Stiefel.

Notons encore que, comme remarqué dans [6], notre décomposition en matrices  $Q$ ,  $U$  est numériquement meilleure que l'élimination de Jordan utilisée dans [5], par exemple.

- (2) Du point de vue implémentation, nous avons programmé un algorithme adapté aux normes  $l_1$  et  $l_\infty$ . Nous nous sommes également limités au cas du problème non-constrait puisque c'est le type de problème que nous aurons à résoudre dans les chapitres suivants.



REFERENCES

---

- [ 1 ] S.K. BERBERIAN, "Functional Analysis and Operator Theory", Springer-Verlag, Berlin, 1974.
- [ 2 ] R.T. ROCKAFELLAR, "Convex Analysis", Princeton University Press, Princeton, N.J., 1970.
- [ 3 ] G. HADLEY, "Linear Programming", Addison-Wesley, Reading Mass., 1965.
- [ 4 ] J. KOWALIK et M.R. OSBORNE, "Methods for Unconstrained Optimization Problems", Elsevier, New York, 1968.
- [ 5 ] E.W. CHENEY, "Introduction to Approximation Theory", McGraw-Hill, New York, 1966.
- [ 6 ] R.H. BARTELS et G.H. GOLUB, "Stable Numerical Methods for Obtaining the Chebyshev Solution to an Overdetermined System of Equations", Comm. ACM.11, 1968, pp. 401-406.
- [ 7 ] D.H. ANDERSON, "Linear Programming and the Calculation of Maximum Norm Approximation", Thèse de Doctorat, Australian National University, avril 1975.
- [ 8 ] E. STIEFEL, "Über Diskrete und Lineare Tschebyscheff-Approximationen", Numer. Math. 1, 1959.

## CHAPITRE II :

UN ALGORITHME DE DESCENTE POUR UNE  
MINIMISATION EN NORME POLYÉDRALE

## § 1. INTRODUCTION ET POSITION DU PROBLEME

---

Dans la première partie, nous avons défini une norme polyédrale et établi un algorithme d'échange pour la résolution d'un problème du type

$$\min_{h \in \mathbb{R}^p} \|f + A h\|_B$$

A présent, nous allons utiliser cet algorithme pour établir un algorithme de descente permettant de résoudre les problèmes du type

$$\min_{x \in \mathbb{R}^p} \|f(x)\|_B \quad (1.1)$$

où  $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$  est de classe  $C^1$ .

Il est intéressant de noter, comme vu au chapitre 1, que le problème (1.1) englobe des problèmes de la forme

$$\min_{x \in \mathbb{R}^p} \max_{i=1, \dots, n} |f_i(x)|$$

$$\min_{x \in \mathbb{R}^p} \sum_{i=1}^n |f_i(x)|$$

étant donné la définition des normes polyédrales.

Hypothèse de travail.

On fera, dans tout ce qui suit, les hypothèses suivantes :

soit  $f : \mathbb{R}^p \rightarrow \mathbb{R}^n$  ( $p < n$ ), de classe  $C^1$  et supposons que

i)  $f$  admet un point minimum local dans l'ensemble

$$R_c = \{x \mid F(x) \leq F(x_0) + c\}$$

où

$$\begin{cases} F(x) = \|f(x)\|_B \\ \text{déf.} \\ c \text{ est une constante positive} \end{cases} ;$$

ii)  $R_c$  est borné ; (1.2)

iii) pour tout  $x, y \in R_c$ ,

$$f(y) = f(x) + A(x)(y - x) + \|y - x\|^2 w(x, y) \quad (1.3)$$

où -  $A(x)$  est la matrice jacobienne de  $f$  au point  $x$ ,

-  $\|\cdot\|$  est une norme sur  $\mathbb{R}^p$

-  $w(x, y)$  est un vecteur dépendant de  $(x, y)$  tel que,  
pour tout  $x, y \in R_c$ ,

$$\|w(x, y)\|_B \leq W ; \quad (1.4)$$

-  $A(x)$  est de rang  $p$ , pour tout  $x \in R_c$ .

Remarque.

Les hypothèses ci-dessus (et particulièrement l'hypothèse (1.2)) peuvent sembler assez fortes. Notons cependant que nos résultats seront essentiellement locaux, ceci permettant de nous limiter à un sous-ensemble de  $R_c$  où nos hypothèses seront vérifiées. Nous pouvons également donner une condition suffisante pour que les hypothèses (1.3) et (1.4) soient satisfaites, à savoir :

Si  $f$  est de classe  $C^2$  et si  $R_0$  est un ensemble convexe et borné, les hypothèses (1.3) et (1.4) sont satisfaites dans  $R_0$ .

Preuve :

$f(x)$  est une fonction de  $\mathbb{R}^p$  dans  $\mathbb{R}^n$ ; elle peut être envisagée comme un vecteur  $(f_1, \dots, f_n)$  où

$$f_i : \mathbb{R}^p \rightarrow \mathbb{R}.$$

En appliquant la formule de Taylor à chaque  $f_i$ , nous obtenons, pour tout  $x, y \in R_0$ ,

$$f_i(y) = f_i(x) + A_i(x) (y - x) + (y - x)^T H_i(\lambda x + (1 - \lambda) y) (y - x)$$

où -  $H_i(z)$  désigne la matrice hessienne au point  $z$  de  $\mathbb{R}^p$ ,  
 -  $\lambda \in [0, 1]$ .

Posant, dès lors,

$$W_i(x, y) = \frac{(y - x)^T H_i(\lambda x + (1 - \lambda) y) (y - x)}{\|y - x\|_2^2},$$

on obtient la relation (1.3) avec

$$|W_i(x, y)| \leq \|H_i(\lambda x + (1 - \lambda) y)\|_2 \leq K_i$$

car l'ensemble  $R_0$  est compact et la matrice hessienne est continue sur ce compact.

Il nous suffit, dès lors, de poser

2.4.

$$(W(x, y))_i = w_i(x, y), \quad i = 1, \dots, n$$

$$\|\cdot\| = \|\cdot\|_2$$

et

$$W = \left( \sum_{i=1}^n K_i^2 \right)^{1/2}$$

pour obtenir le résultat cherché.  $\square$



## § 2. CARACTERISATION D'UN POINT STATIONNAIRE

---

Le critère classique de détermination d'un point minimal est habituellement effectué sur le gradient de la fonction-objectif; malheureusement, le gradient de la fonction-objectif  $\|f(x)\|_B$  est difficile à calculer et n'existe pas partout (par exemple, la norme  $l_\infty$  n'est pas dérivable à l'origine). Dès lors, nous remplacerons cette condition nécessaire d'optimalité par une autre plus adaptée à notre algorithme. A cette fin, introduisons la définition suivante :

Nous dirons que  $x^*$  est un point stationnaire de  $F(x)$  ssi

$$F(x^*) = \lambda^* = \min_{t \in \mathbb{R}^p} \|\phi(x^*) + A(x^*) t\|_B$$

Le problème de minimisation (1.1) peut encore s'écrire

$$\left. \begin{array}{l} \min_{(X, \lambda) \in \mathbb{R}^{n+1}} [0, \dots, 0, 1] \begin{pmatrix} x \\ \lambda \end{pmatrix} \\ \text{sous contraintes} \\ B f(x) \leq \lambda b \end{array} \right\} \quad (2.1)$$

Dans la théorie de la programmation, les conditions nécessaires de Kuhn-Tucker sont bien connues (voir, par exemple, [1, p. 56]). Nous allons vérifier qu'il existe une correspondance biunivoque entre point stationnaire et point vérifiant les conditions de Kuhn-Tucker pour le problème (2.1).

Pour cela, écrivons les conditions de Kuhn-Tucker pour (2.1), c'est-à-dire

$(x^*, \lambda^*)$ , admissible pour (2.1), est une solution optimale si



2.6.

il existe des nombres  $\mu_i \geq 0$ ,  $i = 1, \dots, m$ , tels que

$$(1) \quad \mu_i [\rho_i(B) f(x^*) - \lambda^* b_i] = 0 \quad ; \quad (2.2)$$

$$(2) \quad \sum_{i=1}^m \mu_i b_i = 1 \quad ; \quad (2.3)$$

$$(3) \quad \sum_{i=1}^m \mu_i \rho_i(B) A(x^*) = \theta_p \quad ; \quad (2.4)$$

où (1) sont des conditions de complémentarité,

(2) et (3) sont des conditions nécessaires d'optimalité du premier ordre.

Commençons par établir la correspondance dans le sens

$$P S \rightarrow P K T \quad .$$

Lemme 2.1.

Si  $x^*$  est un point stationnaire de  $F(x)$ ,  $(x^*, \lambda^*)$  satisfait alors les conditions de Kuhn-Tucker (2.2) - (2.4) .

Preuve :

De manière évidente  $(x^*, \lambda^*)$  est admissible pour (2.1) .

D'autre part, le problème  $\min_{t \in \mathbb{R}^p} \|f(x^*) + A(x^*) t\|_B$  peut s'écrire :

2.7.

$$\left. \begin{array}{l} \min_{\lambda, t} [0, \dots, 0, 1] \begin{pmatrix} t \\ \lambda \end{pmatrix} \\ \text{sous les contraintes} \\ [-B \ A, b] \begin{pmatrix} t \\ \lambda \end{pmatrix} \geq B f(x^*) \end{array} \right\} \quad (2.5)$$

son dual ayant la forme

$$\left. \begin{array}{l} \max [w_1, \dots, w_m] [B f(x^*)] \\ \text{sous les contraintes} \\ [w_1, \dots, w_n] [-B \ A, b] = [0, \dots, 0, 1] \\ w_i \geq 0, \quad i = 1, \dots, m \end{array} \right\} \quad (2.7)$$

Désignons par  $w^*$  et  $(t^*, \lambda^*)$  des solutions optimales des problèmes dual et primal respectivement. Par le théorème des écarts complémentaires [cf. Appendice 1.4], nous savons que

$$w_i (-\rho_i(B) A t^* + \lambda^* b_i - \rho_i(B) f(x^*)) = 0 \quad i = 1, \dots, m.$$

Posant, dès lors,

$$\mu_i = w_i, \quad i = 1, \dots, m$$

et puisque, par hypothèse, nous pouvons choisir  $t^* = \theta_p$ , nous obtenons

$$\mu_i (\rho_i(B) f(x^*) - \lambda^* b_i) = 0$$

ce qui est (2.2), et, par l'admissibilité de  $\mu$  pour le problème dual,

$$\mu^T b = 1, \quad ,$$

ce qui est (2.3) , et

$$\mu^T B A(x^*) = \theta_p, \quad ,$$

ce qui est (2.4) .□

Etudions à présent la correspondance

$$P \times K \times T \rightarrow P \times S \quad .$$

Lemme 2.2.

*Si  $[x^*, \lambda^*]$  , admissible pour (2.1) satisfait les conditions de Kuhn-Tucker (2.2) - (2.4) ,  $x^*$  est un point stationnaire de  $F(x)$  .*

Preuve :

Celle-ci résulte immédiatement du fait que les conditions nécessaires de Kuhn-Tucker (2.2) - (2.4) sont les conditions suffisantes (suffisantes car le problème est linéaire) pour que  $(\theta_p, \lambda^*)$  soit solution du problème

$$\min_{t \in \mathbb{R}^p} \|f(x^*) + A(x^*) t\|_B \quad .\square$$

### § 3. UN ALGORITHME DE DESCENTE

---

L'itération générale de notre algorithme de descente aura la forme suivante :  
à chaque itération  $i$ , le vecteur  $t_i$  le long duquel se fait la descente sera déterminé en résolvant le problème

$$\min_{t \in \mathbb{R}^p} \|f(x_i) + A(x_i) t\|_B \quad (3.1)$$

Nous poserons, dès lors,

$$x_{i+1} = x_i + \gamma_i t_i,$$

le paramètre  $\gamma_i$  étant déterminé de manière à assurer la descente de l'algorithme.

Remarques.

---

- (1) Le problème (3.1) peut être résolu grâce à l'algorithme de la première partie.
- (2) Il est traditionnel de choisir  $\gamma_i$  minimisant  $\|f(x_i) + \gamma t_i\|_B$ . Nous allons cependant fournir une autre méthode qui ne réclame pas une minimisation unidimensionnelle et assure la convergence.

Avant de décrire l'algorithme, introduisons quelques définitions.

Nous poserons

$$\psi(x_i, \gamma) = \frac{F(x_i) - F(x_i + \gamma t_i)}{(F(x_i) - \lambda_i)} \quad (3.2)$$

où

$$\lambda_i = \min_{t \in \mathbb{R}^n} \|f(x_i) + A(x_i) t\|_B = \|f(x_i) + A(x_i) t_i\|_B \quad ; \quad (3.3)$$

$$r_i = f(x_i) + A(x_i) t_i \quad ; \quad (3.4)$$

$$\Pi = \{1, \theta, \theta^2, \theta^3, \dots\} \quad 0 < \theta < 1 \quad .$$

Remarquons immédiatement que ces définitions impliquent que

$$F(x_i) \geq \lambda_i \quad . \quad (3.5)$$

Algorithme.

---

Pas 0 : Initialisation.

---

$$x_0, f(x_0), A(x_0), F(x_0), \text{eps}, \delta$$

$$i : = 0$$

Pas 1.

---

Déterminons, par l'algorithme d'échange,  $t_i$  minimisant

$$\|f(x_i) + A(x_i)t\|_B$$


---

Pas 2.

---

Si

$$F(x_i) - \lambda_i \leq \text{eps}$$

alors : STOP , est à l'optimum .

Autrement,

prendre  $\gamma_i$  le plus grand élément de  $\Pi$  tel que

$$\psi(x_i, \gamma_i) \geq \delta \quad (3.6)$$

où  $\delta$  est une constante positive indépendante de  $x_i$  .

Notons que la propriété de descente de l'algorithme est une conséquence de (3.6) puisque, par définition de  $\psi(x_i, \gamma_i)$  , (3.6) équivaut à

$$\frac{F(x_i) - F(x_i + \gamma_i t_i)}{\gamma_i (F(x_i) - \lambda_i)} \geq \delta$$

$$\Leftrightarrow F(x_i + \gamma_i t_i) \leq F(x_i) - \delta \gamma_i (F(x_i) - \lambda_i)$$

et, puisque  $x_i$  n'est pas optimum, nous avons

$$F(x_i) - \lambda_i > \text{eps} > 0$$

Pas 3.

---

Poser

$$x_{i+1} = x_i + \gamma_i t_i$$



$$i : i+1$$

et retourner au Pas 1.

La clé de l'algorithme est la relation (3.6) ; nous allons montrer que, sous nos hypothèses, cette relation est toujours possible. Auparavant, nous aurons toutefois besoin d'un lemme traditionnel pour les algorithmes de descente.

Lemme 3.1.

*Si  $x_i \in R_c$ , alors  $x_{i+1} \in R_c$ .*

Preuve :

De ce qui précède, on tire

$$F(x_{i+1}) < F(x_i) \leq F(x_0) \quad .\square$$

Théorème 3.1.

*Si  $F(x_i) \neq \lambda_i$  et  $0 < \delta < 1$ , il existe alors  $\gamma_i \in \Pi$  tel que la relation (3.6) soit satisfaite.*

Preuve :

Nous allons démontrer que

$$\psi(x_i, \gamma) \geq 1 + o_i(\gamma)$$



où, pour  $x_i$  fixé,

$$\lim_{\gamma \rightarrow 0} O_i(\gamma) = 0$$

En effet, en utilisant (1.3), nous avons

$$f(x_i + \gamma t_i) = f(x_i) + \gamma A_i t_i + \gamma^2 \|t_i\|^2 w(x_i, x_i + \gamma t_i)$$

Utilisant (3.4), on tire

$$f(x_i + \gamma t_i) = f(x_i) + \gamma (r_i - f(x_i)) + \gamma^2 \|t_i\|^2 w(x_i, x_i + \gamma t_i)$$

Puisque  $0 < \gamma < 1$  et

$$\|w(x_i, x_i + \gamma t_i)\|_B \leq W,$$

par (1.4), on a

$$F(x_i + \gamma t_i) \leq (1 - \gamma) F(x_i) + \gamma \lambda_i + \gamma^2 \|t_i\|^2 W \quad (3.7)$$

Portant (3.7) dans (3.2), nous obtenons

$$\psi(x_i, \gamma) \geq \frac{\gamma (F(x_i) - \lambda_i) - \gamma^2 \|t_i\|^2 W}{\gamma (F(x_i) - \lambda_i)}$$

d'où

$$\psi(x_i, \gamma) \geq 1 - O_i(\gamma)$$

où

$$O_i(\gamma) = \frac{\gamma \|t_i\|^2 W}{F(x_i) - \lambda_i}$$

□

Nous allons, à présent, analyser la convergence de l'algorithme.

Avant d'aborder ce problème, précisons quelques notations.

Tout d'abord, la première partie a montré que le vecteur  $[t_i, \lambda_i]$ , solution de

$$\min_{t \in \mathbb{R}^p} \|f(x_i) + A(x_i) t_i\|_B, \quad (3.8)$$

était caractérisé par  $p+1$  lignes de la matrice  $[B A(x_i), -b]$  que nous désignons par  $[B A_i, -b]^{\sigma_j}$  où  $\sigma_j$  était un index parcourant toutes les références possibles formées de  $(p+1)$  lignes de  $[B A_i, -b]$ .

Nous allons modifier quelque peu cette notation, à savoir que, dorénavant, la référence optimale du  $i$ -ième sous-problème linéaire (3.8) sera désignée par  $\sigma_i$ .

Par le théorème (2.1) du chapitre I, nous savons, d'autre part, que  $[B A(x_i), -b]^{\sigma_i}$  est de rang  $p+1$  et qu'il existe des multiplicateurs  $\mu_1^i, \dots, \mu_{p+1}^i \geq 0$  tels que

$$\left. \begin{aligned} [0, \dots, 0, 1] &= \sum_{j \in \sigma_i} \mu_j^i [-\rho_j(B) A_i, b_j] \\ \text{et} \\ \rho_j(B) [f(x_i) + A_i t_i] &= \lambda_i, \quad j \in \sigma_i \end{aligned} \right\} \quad (3.9)$$

Avant de démontrer la convergence, nous aurons besoin d'un lemme technique.

Lemme 3.2.

Il existe une constante  $K > 0$ , indépendante de  $x_i$ , telle que, pour tout  $x_i \in R_0 = \{x \mid F(x) \leq F(x_0)\}$ , on ait

$$\|t_i\|_{\infty} \leq K F(x_i)$$

Preuve :

Par la remarque (3.9), il existe  $\sigma_i$  tel que

$$[B^{\sigma_i} A_i, -b^{\sigma_i}] \begin{pmatrix} t_i \\ \lambda_i \end{pmatrix} = -B^{\sigma_i} f(x_i)$$

D'autre part, comme  $\text{rg} ([B^{\sigma_i} A_i, -b^{\sigma_i}]) = p+1$ , il existe une sous-matrice  $Q(x_i)$  ( $p, p$ ) non singulière telle que

$$Q(x_i) t_i = -\tilde{B}^{\tilde{\sigma}_i} f(x_i) + \lambda_i \tilde{b}^{\tilde{\sigma}_i}$$

Le symbole  $\sim$  indique que nous avons choisi les lignes de  $B^{\sigma_i}$  et  $b^{\sigma_i}$  correspondant à  $Q(x_i)$ .

Nous pouvons, sans perte de généralité, supposer que

$$b_i = 1 \quad i = 1, \dots, m$$

Posons aussi

$$v(x_i) = -\tilde{B}^{\tilde{\sigma}_i} f(x_i) + \lambda_i \tilde{b}^{\tilde{\sigma}_i}$$

Dès lors, comme

$$|\rho_j(B) f(x_i)| \leq \|f(x_i)\|_B \quad j = 1, \dots, m$$

et, par (3.5) ,

$$\|f(x_i)\|_B \geq \lambda_i$$

nous obtenons

$$Q(x_i) t_i = v(x_i)$$

avec

$$|v_j(x_i)| \leq 2 F(x_i) \quad j = 1, \dots, p$$

(3.10)

A présent, par continuité du déterminant, nous pouvons trouver, pour tout  $x_i \in R_0$ , une boule  $B(x_i, \tau_i)$  telle que, pour tout  $x \in B(x_i, \tau_i)$ ,

$$\det(Q_i(x)) \neq 0$$

Soit  $x \in B(x_i, \tau_i)$ .

Deux cas sont possibles :

(1)  $\sigma_i$  est optimale pour  $x$  :

Dès lors, en inversant le système (3.10), on obtient

$$t(x) = Q_i^{-1}(x) v(x)$$

et

$$\|t(x)\|_\infty \leq \|Q_i^{-1}(x)\|_\infty \|v(x)\|_\infty \quad (3.11)$$

(2)  $\sigma_i$  est non-optimale pour  $x$  :

On a, de toute façon,

$$[Q_i(x) \ t(x)]_j \leq v_j(x) \quad j = 1, \dots, p$$

et

$$[-Q_i(x) \ t(x)]_j \leq v_j(x) \quad j = 1, \dots, p$$

car  $t(x)$  doit être admissible pour le système linéaire dérivé de (3.8) .

Dès lors,

$$\|Q_i(x) \ t(x)\|_\infty \leq \|v(x)\|_\infty \leq 2 F(x)$$

Or,

$$\|t(x)\|_\infty = \|Q_i^{-1}(x) \ [Q_i(x) \ t(x)]\|_\infty \leq \|Q_i^{-1}(x)\|_\infty \|Q_i(x) \ t(x)\|_\infty .$$

Donc,

$$\frac{\|t(x)\|_\infty}{\|Q_i^{-1}(x)\|_\infty} \leq 2 F(x)$$

et

$$\|t(x)\|_\infty \leq 2 F(x) \|Q_i^{-1}(x)\|_\infty \quad (3.12)$$

Dès lors, en prenant, pour  $\tau_i$ , la moitié de la distance entre  $x_i$  et le point le plus proche de  $x_i$ , tel que

$$\det(Q_i(x)) = 0$$

et en posant

$$K_i = \sup_{x \in B(x_i, \tau_i)} \|Q_i^{-1}(x)\|_\infty ,$$

on tire de (3.11) et de (3.12)

2.18.

$$\|t(x)\|_{\infty} \leq 2 K_i F(x) \quad (3.13)$$

Nous pouvons toujours choisir les  $x_i$  de telle manière que  $R_0$  soit recouvert par ces boules  $B(x_i, \tau_i)$ .

Comme  $R_0$  est compact, nous pouvons en extraire un sous-recouvrement fini  $B(x_{i_1}, \tau_{i_1}), \dots, B(x_{i_p}, \tau_{i_p})$ .

Posant

$$K = \max_{j=1, \dots, p} K_{i_j},$$

on obtient le résultat recherché, à savoir

$$\|t(x)\|_{\infty} \leq 2 K F(x) \quad .\square \quad (3.14)$$

Nous sommes, à présent, capables de démontrer la convergence de l'algorithme vers un point stationnaire de  $F(x)$ , au sens du paragraphe 2.

Théorème 3.2 : Convergence.

*Si l'algorithme ne se termine pas en un nombre fini d'étapes, alors*  

$$(F(x_i) - \lambda_i) \xrightarrow{i \rightarrow \infty} 0.$$

Preuve :

(1) La suite  $F(x_i)$  converge :

En effet, par le choix de  $x_{i+1}$  (Pas 2 de l'algorithme), on a

$$F(x_{i+1}) \leq F(x_i) - \delta \gamma_i (F(x_i) - \lambda_i) \quad (3.15)$$

La suite  $(F(x_i))_i$  est donc décroissante et bornée inférieurement par zéro, donc convergente.

(2) Démontrons la convergence en deux étapes :

1) Il existe  $\gamma > 0$  tel que  $\gamma_i > \gamma$ , pour tout  $i$ . Réarrangeant (3.15), on obtient

$$F(x_i) - \lambda_i \leq \frac{1}{\gamma \delta} (F(x_i) - F(x_{i+1}))$$

et, par l'assertion ci-dessus,  $(F(x_i) - \lambda_i) \rightarrow 0$ .

2) Supposons que

$$\hat{\gamma} = \inf \{\gamma_i\} = 0$$

Nous pouvons, dès lors, trouver une suite  $\gamma_i^0$  telle que

$$\psi(x_i, \gamma_i^0) < \delta$$

(par exemple  $\gamma_i^0 = \gamma_i / \theta$ ,  $\gamma_i \neq 1$ ).

Dès lors,

$$1 - \frac{\gamma_i^0 \|t_i\|^2 W}{F(x_i) - h_i} < \delta$$

D'où

$$F(x_i) - \lambda_i < \gamma_i^0 \|t_i\|^2 / (1 - \delta)$$

$$< 2 \gamma_i^0 K' F(x_0) W / (1 - \delta) \quad (3.16)$$



où

$$K' = K k^2 ,$$

$K$  est la constante de la relation (3.14) et  $k$  une constante telle que

$$\| \cdot \| \leq K \| \cdot \|_{\infty}$$

Comme  $\gamma_i^0 \rightarrow 0$ , on obtient le résultat recherché.  $\square$

### Corollaire 3.2.

Si une sous-suite  $(x_{\pi i})$  de  $(x_i)$  converge, son point limite est un point stationnaire de  $F(x)$ .

#### Preuve :

Un théorème bien connu d'analyse affirme que

si  $C(x, t)$  est une fonction continue en  $x$  et  $t$ ,

si  $t$  appartient à un compact  $K$ ,

alors la fonction

$$G(x) = \min_{t \in K} C(x, t)$$

est continue en  $x$ .

Posant, dès lors,

$$C'(x, t) = \| f(x) + A(x) t \|_B$$

et comme, par le Lemme (3.2) ,

$$\|t(x)\|_{\infty} \leq 2 K F(x_0) ,$$

la fonction

$$G(x) = \min_{\|t\|_{\infty} \leq 2KF(x_0)} C'(x, t)$$

est continue en  $x$  .

Dès lors, si  $x_{\pi i} \rightarrow x^*$  , nous avons

$$\lambda_{\pi i} = G(x_{\pi i}) \rightarrow \lambda^* = G(x^*)$$

D'autre part, le théorème (3.2) donne

$$F(x_{\pi i}) - \lambda_{\pi i} \rightarrow 0$$

d'où l'on conclut

$$F(x^*) = \lambda^* \quad .\square$$

Naturellement, la question suivante se pose immédiatement après la question de convergence : quelle est la vitesse de convergence de l'algorithme ?

Nous aborderons ce problème dans un cadre moins général que celui dans lequel nous avons travaillé jusqu'à présent.

Nous allons également voir que ces nouvelles conditions sont suffisantes pour assurer la convergence de la suite  $(x_i)$  vers un point  $x$  , qui, par le corollaire (3.2) , sera un point stationnaire de  $F(x)$  .

## § 4. HYPOTHESES SUFFISANTES DE CONVERGENCE

---

### Définition.

---

Nous dirons que la règle des multiplicateurs est satisfaite sur le compact  $S \subset R_0$  ssi

(1) il existe une référence optimale  $\sigma_i$  identique pour tout  $x_i \in R$ ;

(2) pour tout  $j \in \sigma_i$  et tout  $x_i \in S$ ,

$$\mu_j^i \geq \hat{\mu} > 0$$

Nous allons établir, grâce à cette règle, une nouvelle borne sur  $t_i$ ; ce sera l'objet du lemme suivant qui fournit également une condition nécessaire pour que la règle des multiplicateurs soit satisfaite. Nous verrons également comment exploiter numériquement cette condition.

### Lemme 4.1.

---

Si la règle des multiplicateurs est satisfaite sur  $S$  compact  $\subset R_0$ , alors, pour tout  $x_i \in S$ ,

$$\|t_i\|_{\infty} \leq K (F(x_i) - \lambda_i) / \hat{\mu}$$

Preuve :

Posons

2.23.

$$a = B^{\sigma_i} f(x_i) - \lambda_i b^{\sigma_i} \quad (\text{N.B. } b_i = 1, \forall i = 1, \dots, m)$$

$$\Delta_i = F(x_i) - \lambda_i$$

$$a = a_+ - a_- \quad \begin{cases} a_+ = \max(a, 0) \\ a_- = \max(-a, 0) \end{cases} \quad \text{ou} \quad (a_+)_j = \begin{cases} 0 & \text{si } a_j \leq 0 \\ a_j & \text{si } a_j > 0 \end{cases}$$

Nous allons prouver que

$$\begin{cases} e^t a_+ \leq \Delta_i / \hat{\mu} \end{cases} \quad (4.1)$$

$$\begin{cases} e^t a_- \leq \Delta_i / \hat{\mu} \end{cases} \quad (4.2)$$

où

$$e^t = (1, \dots, 1)$$

En admettant ces deux relations, nous obtenons donc

$$\max_{j \in \sigma_i} |a_j| \leq e^t a_+ \leq \Delta_i / \hat{\mu}$$

ou bien

$$\max_{j \in \sigma_i} |a_j| \leq e^t a_- \leq \Delta_i / \hat{\mu}$$

c'est-à-dire

$$\|a\|_{\infty} \leq \Delta_i / \hat{\mu}$$

Or,

$$[B^{\sigma_i} A_i, -b^{\sigma_i}] t_i = a$$

Dès lors, comme, par la règle des multiplicateurs  $[B^{\sigma_i} A_i, -b^{\sigma_i}]$  est non-singulière sur  $S$  compact, le même raisonnement que pour le lemme (3.2) permet d'affirmer qu'il existe  $K > 0$  tel que, pour tout  $x \in S$ ,

$$\|t(x)\|_{\infty} \leq K \|a\|_{\infty} \leq K (F(x) - \lambda) / \hat{\mu}$$

Pour terminer la démonstration, il suffit de prouver (4.1) et (4.2).

$$(1) \quad \frac{e^t a_+}{\hat{\mu}} \leq \Delta_i / \hat{\mu} :$$

Nous avons  $0 \leq \Delta_i$  et

$$|(B^{\sigma_i} f(x_i))_j| \leq F(x_i) \quad j = 1, \dots, p+1$$

$$\Rightarrow a_+ \leq \Delta_i e$$

$$\Rightarrow \hat{\mu} a_+ \leq \hat{\mu} \Delta_i e \leq \mu^i \Delta_i \quad \text{ou} \quad \mu^i = (\mu_1^i, \dots, \mu_{p+1}^i)$$

$$\Rightarrow \hat{\mu} e^t a_+ \leq e^t \mu^i \Delta_i$$

et, par (3.9),

$$e^t \mu^i = 1$$

$$(2) \quad \frac{e^t a_-}{\hat{\mu}} \leq \Delta_i / \hat{\mu} :$$

car

$$\hat{\mu} e^t a_- \leq \mu_i^t a_-$$

Or, par définition de  $\sigma_i$ , (3.9) implique que

$$a_{\mu}^t i = -\mu_i^t B^{\sigma i} A_i t_i = 0$$

Dès lors,

$$\hat{\mu} e^t a_{-} \leq a_{-}^t \mu^i = a_{+}^t \mu^i$$

$$\leq e^t \mu^i \Delta_i$$

$$= \Delta_i$$

□

Ce lemme fournit une manière pour déterminer si la règle des multiplicateurs est satisfaite. Il suffit, à chaque itération, de calculer  $\|t_i\|_{\infty} / (F(x_i) - h_i)$ .

Si cette quantité s'accroît lorsque  $i$  augmente, il est probable que la règle des multiplicateurs n'est pas satisfaite. Dès lors, le théorème de convergence quadratique que nous allons voir peut ne pas être satisfait; d'où une convergence qui peut devenir linéaire, ce qui, numériquement, peut impliquer un très grand nombre d'itérations avant de vérifier le test d'arrêt.

La règle des multiplicateurs peut également simplifier très fort la détermination du paramètre  $\gamma_i$ , puisque cette règle assure qu'il y a moyen de prendre  $\gamma_i = 1$  après un certain nombre d'itérations.

Lemme 4.2.

*S'il existe  $i'$  tel que, pour  $i > i'$ ,  $x_i \in S$  compact et  $S \subset R_0$  et si la règle des multiplicateurs est satisfaite sur  $S$ , alors il existe  $i_0$  tel que, pour tout  $i > i_0$ ,  $\gamma_i = 1$ .*

Preuve :

Par le lemme (4.1) et le théorème (3.2), nous avons

$$\|t_i\|_{\infty} \xrightarrow{i \rightarrow \infty} 0$$

Dès lors, il existe  $i'_0$  tel que

$$i > i'_0 \Rightarrow x_i + t_i \in R_0$$

En utilisant les résultats des lemmes (3.1) et (4.1), nous obtenons

$$\psi(x_i, 1) \geq 1 - \frac{\|t_i\|^2 W}{F(x_i) - \lambda_i}$$

et

$$\psi(x_i, 1) \geq 1 - (K/\hat{\mu})^2 W(F(x_i) - \lambda_i)$$

Par le théorème (3.2),

$$F(x_i) - \lambda_i \rightarrow 0$$

Dès lors, il existe  $i' > i'_0$  tel que, pour tout  $i > i'$ ,

$$\psi(x_i, 1) > \delta \qquad 0 < \delta < 1 \quad .\square$$

A ce stade-ci, nous sommes capables d'assurer la convergence de la suite  $(x_i)$  et, par le corollaire (3.2), nous pouvons, dès lors, assurer que le point limite de la suite  $(x_i)$  est un point stationnaire de  $F(x)$ .



Théorème 4.1.

Sous les hypothèses du Lemme (4.2), la suite  $(x_i)$  est de Cauchy.

Preuve :

Posons

$$S_{\infty} = \sum_{i=i_{\sigma}}^{\infty} \|t_i\|_{\infty}$$

où  $i_{\sigma}$  est l'indice  $i_0$  du lemme (4.2).

Par le lemme (4.1), nous avons

$$\begin{aligned} S_{\infty} &\leq (K/\hat{\mu}) \sum_{i=i_{\sigma}}^{\infty} (F(x_i) - \lambda_i) \\ &\leq (K/\hat{\mu} \delta) \sum_{i=i_{\sigma}}^{\infty} (F(x_i) - F(x_{i+1})) \quad \text{par (3.15)} \end{aligned}$$

Dès lors,

$$S_{\infty} \leq (K/\hat{\mu} \delta) F(x_{i_{\sigma}})$$

et est donc bornée.

Posons

$$S_r = \sum_{i=r}^{\infty} \|t_i\|_{\infty}$$

Soit  $\varepsilon > 0$ ;  $p, q \geq i_{\sigma}$ .

Nous avons

$$|S_p - S_q| = \sum_{i=q+1}^p \|t_i\|_\infty$$

Ceci implique qu'il existe  $N(\epsilon)$  tel que

$$p, q > N(\epsilon) \Rightarrow |S_p - S_q| \leq \epsilon$$

Dès lors,

$$\|x_p - x_q\|_\infty \leq \sum_{i=q+1}^p \|t_i\|_\infty = |S_p - S_q|$$

Donc, pour tout  $\epsilon > 0$ , il existe  $N(\epsilon)$  tel que

$$p, q > N(\epsilon) \Rightarrow \|x_p - x_q\|_\infty < \epsilon$$

Nous sommes à présent en mesure d'analyser la vitesse de convergence de l'algorithme; auparavant, démontrons un lemme. Ce lemme indique simplement que, si  $\|t_i\| \rightarrow 0$  et si  $j$  est un indice tel que la  $j$ -ième contrainte du problème linéaire

$$\min_{t \in \mathbb{R}^p} \|f(x^*) + A(x^*) t\|_B$$

est saturée, c'est-à-dire

$$\rho_j(B) f(x^*) + \rho_j(B) A(x^*) t(x^*) = \lambda^* b_j,$$

alors, la  $j$ -ième contrainte du problème linéaire

$$\min_{t \in \mathbb{R}^p} \|f(x_i) + A(x_i) t\|_B$$

sera aussi saturée pour  $i$  assez grand.

Pour cela, introduisons deux définitions :

$$\Sigma_i = \{j \mid \rho_j(B) [f(x_i) + A_i t_i] = \lambda_i\}$$

$$\Sigma^* = \{j \mid \rho_j(B) f(x^*) = F(x^*) = \lambda^*\}$$

Lemme 4.3.

Si

$$\|t_i\| \rightarrow 0$$

et si

$$\sum_{i=1}^{\infty} \gamma_i t_i$$

converge, alors il existe un index  $I$  tel que, pour tout  $i > I$ ,  $\Sigma_i \subset \Sigma^*$ .

Preuve :

Soit  $k \in \Sigma^*$  et  $j \notin \Sigma^*$ .

Dès lors, il existe  $\eta > 0$  tel que

$$\rho_k(B) f(x^*) > \rho_j(B) f(x^*) + \eta$$

Comme  $\sum \gamma_i t_i$  converge, nous avons  $\gamma_i t_i \rightarrow 0$  et, par continuité de  $f(x)$ , il existe un index  $v^j$  tel que, pour tout  $i > v^j$ ,

$$\rho_k(B) f(x_i) > \rho_j(B) f(x^*) + \varepsilon \quad \varepsilon > 0$$

Puisque  $\|t_i\| \rightarrow 0$  lorsque  $i \rightarrow \infty$ , il existe un index  $I^j \geq v^j$  tel que, pour tout  $i > I^j$ ,

$$\rho_k(B) [f(x_i) + A_i t_i] > \rho_j(B) [f(x_i) + A_i t_i]$$

Comme

$$\lambda_i \geq \rho_k(B) [f(x_i) + A_i t_i] > \rho_j(B) [f(x_i) + A_i t_i],$$

on a

$$j \notin \Sigma_i$$

Par conséquent,

$$\Sigma_i \subset \Sigma^* \quad , \quad \forall i > I = \max_{j \in \Sigma^*} I^j \quad .\square$$

Théorème 4.2 : Vitesse de Convergence.

*Si, pour  $i$  assez grand,  $x_i \in S$  compact et si la règle des multiplicateurs est satisfaite sur  $S$ , alors la vitesse de convergence est quadratique.*

Preuve :

Nous allons montrer qu'il existe un index  $I$  tel que, pour tout  $i > I$ ,

$$\|x_{i+1} - x^*\|_\infty \leq K \|x_i - x^*\|^2$$

où  $K$  est une constante positive et

$$x^* = \lim_{i \rightarrow \infty} x_i$$

Par le lemme (4.2) , il existe un index  $I$  tel que, pour tout  $i > I$  ,

$$\gamma_i = 1$$

Par la première partie du théorème (4.1) , on a que

$$\sum_{i=1}^{\infty} \|t_i\|_{\infty} \quad \text{converge}$$

et, par le lemme (4.1) ,

$$\|t_i\|_{\infty} \xrightarrow{i \rightarrow \infty} 0$$

Nous sommes donc dans les conditions d'application du lemme (4.3) .

Il existe un index  $I' > I$  tel que  $k \in \sigma_i$  implique

$$k \in \Sigma^* \Rightarrow \rho_k(B) f(x^*) = \lambda^* \quad (4.3)$$

En utilisant (1.3) , nous pouvons écrire

$$\rho_k(B) f(x^*) = \rho_k(B) [f(x_i) + A_i(x^* - x_i)] + \|x_i - x^*\|^2 \tilde{w}_k \quad (4.4)$$

où

$$\tilde{w}_k = \rho_k(B) w(x_i, x^*)$$

La solution du  $i$ -ième sous-problème linéaire satisfait

$$\lambda_i = \rho_k(B) [f(x_i) + A_i t_i] \quad k \in \sigma_i \quad (4.5)$$

En soustrayant (4.4) de (4.5), nous obtenons

$$\rho_k(B) A_i (x_i + t_i - x^*) - \lambda_i + \rho_k(B) f(x^*) = \|x^* - x_i\|^2 \tilde{w}_k$$

et, en utilisant (4.3),

$$\rho_k(B) A_i (x_i + t_i - x^*) - \lambda_i + \lambda^* = \|x^* - x_i\|^2 \tilde{w}_k \quad (4.6)$$

ou encore

$$[B^{\sigma_i}, -b^{\sigma_i}] \begin{pmatrix} x_{i+1} - x^* \\ \lambda^* - \lambda_i \end{pmatrix} = \|x^* - x_i\|^2 \tilde{w}(x_i, x^*)$$

Comme  $[B^{\sigma_i} A_i, -b^{\sigma_i}]$  est non-singulière,

$$\|x_{i+1} - x^*\|_{\infty} \leq \|[B^{\sigma_i} A_i, -b^{\sigma_i}]^{-1}\|_{\infty} k \|x^* - x_i\|_{\infty}^2$$

où

$$k = K_1^2 W K_2$$

avec  $K_1$  et  $K_2$  des constantes positives telles que

$$\|\cdot\| \leq K_1 \|\cdot\|_{\infty}$$

$$\|\cdot\|_B \leq K_2 \|\cdot\|_{\infty}$$

ce qui est toujours possible par l'équivalence des normes sur  $\mathbb{R}^n$ .

A présent, par un raisonnement identique à celui du lemme (3.2) nous permet de borner uniformément l'expression  $\|[B^{\sigma_i} A_i, -b^{\sigma_i}]^{-1}\|_{\infty}$  par une constante  $K_3$ .

Nous obtenons donc

$$\|x_{i+1} - x^*\|_{\infty} \leq K_3 k \|x^* - x_i\|_{\infty}^2,$$

ce qui démontre la convergence quadratique.  $\square$



## CONCLUSIONS

---

- (1) L'algorithme est beaucoup plus puissant si la règle des multiplicateurs est satisfaite. Cette règle, pratiquement jamais vérifiable à priori, nous suggère de choisir, comme référence de départ lors de la résolution du  $i$ -ième sous-problème linéaire, la référence optimale précédente; si cette manière de procéder conduit à une référence initiale de rang inférieur à  $p$ , le coût d'une factorisation de Gauss sur la matrice  $B^{\sigma_i} A_{i+1}$  est négligeable étant donné le nombre de sous-problèmes linéaires à résoudre si la convergence est linéaire.
- (2) Cette même règle nous suggère aussi de calculer le quotient  $\|t_i\|_{\infty} / F(x_i) - \lambda_i$  afin de calculer s'il dépasse un paramètre arbitrairement grand. Il peut, en effet, être plus intéressant de se limiter à l'approximation obtenue que d'itérer un grand nombre de fois pour gagner des décimales supplémentaires.
- (3) Soulignons encore l'importance d'une hypothèse qui peut être mise en défaut, c'est-à-dire le fait que le rang de la matrice jacobienne soit de rang maximal. Outre le fait que cette condition doit être vérifiée en tout point  $x_i$  de la suite construite pour pouvoir appliquer l'algorithme du chapitre I au sous-problème linéaire

$$\min_{t \in \mathbb{R}^p} \|f(x_i) + A_i t\|_B,$$

cette condition doit aussi être vérifiée à l'optimum.

En effet, prenons un exemple simple :

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto (x^2 + 1)^{1/2}$$

Dès lors,

$$f'(x) = x / (x^2 + 1)^{1/2}$$

implique que

$$\operatorname{rg} (f'(x)) = 1$$

dans  $[-1, 1]$  sauf au point minimum 0 où

$$\operatorname{rg} (f'(x)) = 0$$

On a aussi

$$\min_t |f(x) + A(x) \cdot t| = \min_t |(x^2 + 1)^{1/2} + x \cdot t / (x^2 + 1)^{1/2}|$$

Ce minimum est atteint en

$$t = -(x^2 + 1)^{1/2} / x$$

et vaut 0.

On remarque, dès lors que  $|t_i|$  est non-borné et que, pour tout  $x_i \in [-1, 1] \setminus \{0\}$ ,

$$\lambda_i = 0$$

De plus,

$$F(x_i) - \lambda_i = F(x_i) \rightarrow 1 \quad \text{lorsque} \quad x_i \rightarrow 0$$

Ceci montre que le test d'arrêt ne sera jamais vérifié, ce qui est

normal puisque la fonction

$$\min_t \|f(x) + A(x) t\|$$

n'est pas continue en  $x$ .

Le chapitre III donnera une solution à ce type de problème.

REFERENCES

---

- [ 1 ] M.D. CANON, "Theory of Optimal Control and Mathematical Programming", McGraw-Hill, New York, 1970.
- [ 2 ] J. KOWALIK et M.R. OSBORNE, "Methods for Unconstrained Optimization Problems", Elsevier, New York, 1973.
- [ 3 ] D.H. ANDERSON, "Linear Programming and the Calculation of Maximum Norm Approximation", Thèse de Doctorat, Australian National University, avril 1975.
- [ 4 ] G. HADLEY "Linear Programming", Addison-Wesley, 1962.
- [ 5 ] W.I. ZANGWILL, "Non Linear Programming - A Unified Approach", Prentice-Hall, London, 1969.
- [ 6 ] M.R. OSBORNE et G.A. WATSON, "An Algorithm for Minimax Approximation in the Non-Linear Case", Computer J. 12, 1969, pp. 64-69.

### CHAPITRE III :

UN ALGORITHME DE TYPE LEVENBERG POUR  
UNE MINIMISATION EN NORME POLYÉDRALE

## § 1. INTRODUCTION

---

L'algorithme précédent visant à résoudre les problèmes du type

$$\min \|f(x)\|_B$$

présuppose que la matrice jacobienne est de rang plein sur un compact donné et, en particulier, au point optimal.

Malheureusement, cette condition n'est pas toujours remplie. De plus, un mauvais conditionnement de cette matrice peut conduire à une perte numérique de rang et, éventuellement, à des erreurs numériques importantes.

Pour résoudre ces problèmes, les algorithmes de Levenberg ajoutent une correction à la matrice jacobienne afin d'obtenir une matrice augmentée bien conditionnée.

A chaque itération, on ne résout donc plus le sous-problème linéaire

$$\min_{t \in \mathbb{R}^p} \|f(x) + A(x) t\|_B$$

mais bien

$$\min_{t \in \mathbb{R}^p} \left\| \begin{pmatrix} f(x) \\ 0 \end{pmatrix} + \begin{pmatrix} A(x) \\ C(x) \end{pmatrix} t \right\|_{B^*}$$

où  $\|\cdot\|_{B^*}$  est une norme polyédrale qui sera précisée ultérieurement et  $C(x)$  est une matrice carrée de rang  $p$  que l'on choisit la mieux conditionnée possible (orthogonale, par exemple).

Nous aurons besoin d'un lemme fondamental.

Lemme 1.1.

Si  $B_1 : \mathbb{R}^n \rightarrow \mathbb{R}^m$  et  $B_2 : \mathbb{R}^q \rightarrow \mathbb{R}^s$  sont des matrices définissant des normes polyédrales sur  $\mathbb{R}^n$  et  $\mathbb{R}^q$  respectivement, alors

$$\lambda = \max \{ \|v_1\|_{B_1}, \|v_2\|_{B_2} \}$$

est une norme polyédrale sur  $\mathbb{R}^n \oplus \mathbb{R}^q$  définie par

$$\left\| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\|_B = \min \left\{ \lambda, \left\| \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\| \leq \lambda \begin{pmatrix} e_m \\ e_s \end{pmatrix} \right\}$$

où

$$e_m^t = \underbrace{(1, \dots, 1)}_m$$

La démonstration est évidente.  $\square$



## § 2. BASES POUR LE DEVELOPPEMENT D'UN ALGORITHME

---

Le problème de minimisation que nous désirons résoudre est donc le suivant :

$$\min_x \|f(x)\|_{B_1}$$

où

$\|\cdot\|_{B_1}$  est une norme polyédrale définie par la matrice  $B_1$  ;

(2.1)

$f : \mathbb{R}^p \rightarrow \mathbb{R}^n$  est de classe  $C^1$  ( $n > p$ ).

Nous ajouterons les hypothèses suivantes :

(1) Etant donné une approximation initiale  $x_0$ , l'ensemble

$$R_C = \{x \mid \|f(x)\|_{B_1} \leq \|f(x_0)\|_{B_1} + c\} \quad (2.2)$$

est borné.

(2) Pour tout  $x, y \in R_C$ ,

$$f(y) = f(x) + A(x)(y - x) + \|y - x\|^2 w(x, y) \quad (2.3)$$

où -  $A(x)$  est la matrice jacobienne au point  $x$  ;

-  $w(x, y)$  est un vecteur indépendant de  $x$  et  $y$  tel que, pour tout  $x, y \in R_C$ ,

$$\|w(x, y)\| \leq W ;$$

-  $\|\cdot\|$  est une norme sur  $\mathbb{R}^p$ .

Remarquons que, contrairement aux hypothèses du chapitre II, aucune supposition n'est faite sur le rang de la matrice jacobienne, qui peut être quelconque.

L'idée fondamentale de l'algorithme est de déterminer la correction  $t_i$  au point  $x_i$  de la manière suivante.

Posant

$$H(x, t, \gamma) = \max \{ \|f(x) + A(x) t\|_{B_1}, \gamma \|C(x) t\|_{B_2} \} \quad (2.5)$$

où  $\gamma$  est un réel positif et  $C(x)$  une matrice  $(p, p)$  de rang plein, et continue par rapport à  $x$ , pour tout  $x \in R_c$ , nous déterminons  $t_i$  en résolvant le problème linéaire

$$\min_{t \in \mathbb{R}^p} H(x, t, \gamma) \quad (2.6)$$

Grâce au lemme (1.1), ce problème peut être résolu en utilisant l'algorithme du chapitre I.

Nous supposons, de plus, que, pour tout  $x \in R_c$ ,

$$\left. \begin{aligned} & \max_{\|t\|_{B_2}=1} \|C(x) t\|_{B_2} = 1 \\ & \text{et} \\ & \min_{\|t\|_{B_2}=1} \|C(x) t\|_{B_2} \geq 1/\beta \end{aligned} \right\} \quad (2.7)$$

notons encore

$$\|f(x)\|_{B_1} = F(x)$$

3.5.

et

$$H(x, t(x, \gamma), \gamma) = h(x, \gamma) \quad (2.8)$$

où  $t(x, \gamma)$  est la solution de (2.6) .

Nous allons, à présent, montrer que le paramètre  $\gamma$  contrôle la taille du vecteur  $t_i$  , à savoir

Lemme 2.1.

Si  $t_i$  minimise  $H(x_i, t, \gamma)$  , alors

$$\|t_i\|_{B_2} \leq \beta F(x_i) / \gamma \quad \gamma \neq 0 .$$

Preuve :

Posons

$$t = \theta_p ,$$

Nous avons alors

$$h(x_i, \gamma) = \min_{t \in \mathbb{R}^p} H(x_i, t, \gamma) \leq F(x_i) = \|f(x_i)\|_{B_1} .$$

De plus, par (2.8) et (2.5) ,

$$h(x_i, \gamma) \geq \gamma \|C_i t_i\|_{B_2}$$

et, par (2.7) ,

$$h(x_i, \gamma) \geq \gamma \|t_i\|_{B_2} / \beta \quad .\square$$

Nous avons défini, au chapitre précédent, la notion de point stationnaire de  $F(x)$ .

Comme au chapitre II, nous obtenons encore une caractérisation très simple d'un point stationnaire, plus précisément :

Théorème 2.1.

Si, pour une valeur finie de  $\gamma$ ,

$$F(x^*) = h(x^*, \gamma) \quad , \quad (2.10)$$

alors  $x^*$  est un point stationnaire de  $\|f(x)\|_{B_1}$  au sens du chapitre II, p. 2.5.

Inversément, si  $x$  est un point stationnaire de  $\|f(x)\|_{B_1}$ , alors (2.10) est satisfaite pour tout  $\gamma \in \mathbb{R}^+$ .

Preuve :

(1) Nécessité :

Démontrons, par contraposition, que, si  $x$  n'est pas point stationnaire, il existe  $\varepsilon$ ,  $\xi > 0$ , indépendants de  $\gamma$ , tels que

$$F(x) - h(x^*, \gamma) \geq \varepsilon \tau \quad (2.11)$$

où

$$\tau = \min (1/2, \xi / \gamma)$$

En effet, si  $x$  n'est pas point stationnaire de  $F(x)$ , il existe

3.7.

$t^*$  et  $\varepsilon > 0$  tels que

$$\min_t \|f + A t\|_{B_1} = \|f + A t^*\|_{B_1} = F(x) - \varepsilon \quad (2.12)$$

De plus, pour tout  $\tau' > 0$ ,

$$\begin{aligned} \|f(x) + \tau' A(x) t^*\|_{B_1} &= \|\tau' (f(x) + A(x) t^*) + (1 - \tau') f(x)\|_{B_1} \\ &\leq \tau' (F(x) - \varepsilon) + (1 - \tau') F(x) \\ &= F(x) - \varepsilon \tau' \end{aligned} \quad (2.13)$$

Nous pouvons toujours choisir  $\xi$  tel que

$$\|\xi \subset t^*\|_{B_2} \leq F(x) - \varepsilon / 2$$

Soit  $\tau$  tel que  $\tau$  satisfait (2.13) et

$$\tau \leq \min(1/2, \xi / \gamma)$$

On en conclut que

$$\|\gamma \subset \tau t^*\|_{B_2} \leq F(x) - \varepsilon / 2 \quad (2.14)$$

De (2.14) et (2.13), on tire

$$H(x, \tau t^*, \gamma) \leq F(x) - \varepsilon \tau$$

Comme

$$h(x, \gamma) \leq H(x, \tau t^*, \gamma)$$

nous obtenons alors

$$F(x) - h(x, \gamma) \geq \varepsilon \tau$$

(2) Suffisance :

Soit  $t^*$  minimisant  $H(x^*, t, \gamma)$ .

Dès lors,

$$F(x^*) \geq h(x^*, \gamma) \geq \|f(x^*) + A(x^*) t^*\|_{B_1}$$

$$= \min_t \|f(x^*) + A(x^*) t\|_{B_1}$$

$$= F(x^*)$$

Cette dernière égalité est possible car  $x^*$  est un point stationnaire de  $F(x)$ .  $\square$

Remarques.

(1) La définition de la norme sur  $\mathbb{R}^p \oplus \mathbb{R}^q$  est fondamentale.

Supposons que nous prenons  $p = q = 1$  et  $f(x) = x$ , pour tout  $x \in \mathbb{R}$ , sur  $\mathbb{R}$ , la norme  $l_1$  et sur  $\mathbb{R}^2$ , la norme  $l_1$ .

Alors,

$$\left\| \begin{pmatrix} f(x) \\ 0 \end{pmatrix} + \begin{pmatrix} A(x) \\ \gamma \end{pmatrix} t \right\|_1 = |x+t| + |t| \gamma$$

Soit  $x = 1$  et  $\gamma > 1$ ; on a, pour tout  $t = 0$ ,

$$\min_t (|1+t| + |t| \gamma) = 1$$

Cependant, 1 n'est pas un point stationnaire de  $f(x)$  puisque

$$\min_t |1+t| = 0 \neq F(1) = 1$$

- (2) Par le chapitre I, nous savons que la solution optimale du problème (2.6) est caractérisée par une référence formée de  $p+1$  lignes de la matrice

$$\begin{pmatrix} B_1 & A_i \\ B_2 & C_i \end{pmatrix}$$

Nous désignerons par  $\sigma_i^{(1)}$  l'ensemble comprenant les lignes de  $B_1$  entrant dans la référence optimale et par  $\sigma_i^{(2)}$  l'ensemble comprenant les lignes de  $B_2$ .

Nous pouvons donc écrire

$$\begin{pmatrix} -B_1^{\sigma_i^{(1)}} A_i, e^{\sigma_i^{(1)}} \\ -\gamma B_2^{\sigma_i^{(2)}} C_i, e^{\sigma_i^{(2)}} \end{pmatrix} \begin{pmatrix} t_i \\ h(x_i, \gamma) \end{pmatrix} = \begin{pmatrix} B_1^{\sigma_i^{(1)}} f(x_i) \\ 0 \end{pmatrix} \quad (2.15)$$

tandis que la relation (2.17) du chapitre I s'écrit

$$\begin{aligned} [0, \dots, 0, 1] = & \sum_{k \in \sigma_i^{(1)}} \mu_k^{(i)} [-\rho_k(B_1) A_i, 1] \\ & + \sum_{k \in \sigma_i^{(2)}} v_k^{(i)} [-\gamma \rho_k(B_2) C_i, 1] \end{aligned} \quad (2.16)$$

avec  $\mu_k^{(i)}, v_k^{(i)} \geq 0$  pour tout  $k$ .



### § 3. ALGORITHME

---

#### 3.1. Idée Intuitive.

---

Nous allons construire une suite de points  $(x_i)$  satisfaisant deux conditions :

- (1)  $F(x_i) < F(x_{i-1})$  (algorithme de descente) ;
- (2)  $F(x_i) - h(x_i, \lambda_i) \xrightarrow{i \rightarrow \infty} 0$  (convergence vers un point stationnaire)

Afin d'assurer ces deux conditions, nous utiliserons une fonction

$$\phi(x, t, \gamma) = \frac{F(x) - F(x+t)}{F(x) - h(x, \gamma)} \quad (3.1)$$

Nous choisirons un scalaire  $b$  tel que  $0 < b < 1/2$  et essaierons de déterminer à chaque étape, un  $\gamma_i$  tel que

$$b \leq \phi(x_i, t_i, \gamma_i) \quad (3.2)$$

soit satisfaite.

Remarquons que, par (3.1) , (3.2) peut s'écrire

$$F(x_i + t_i) \leq F(x_i) - b (F(x_i) - h(x_i, \gamma_i))$$

Vu que

$$F(x_i) \geq h(x_i, \gamma_i) ,$$

la condition (3.2) exprime donc le caractère de descente de l'algorithme.

Nous poserons, dès lors,

$$x_{i+1} = x_i + t_i$$

Nous allons, à présent, montrer que, sous les hypothèses faites, la condition (3.2) peut être satisfaite.

### 3.2. Convergence.

---

#### Théorème 3.1.

---

*Si  $x_i \in R_C$  et si  $x_i$  n'est pas un point stationnaire de  $F(x)$ , il existe un scalaire  $\gamma_i$  satisfaisant (3.2).*

Preuve :

Par le lemme (2.1) et la continuité de  $F$ , nous pouvons toujours prendre  $\gamma$  assez grand pour que  $x_i + t_i \in R_C$ .

Dès lors, par (2.3),

3.12.

$$\begin{aligned}
 f(x_i + t_i) &= f(x_i) + A_i t_i + \|t_i\|^2 w(x_i, t_i) \\
 \Rightarrow F(x_i + t_i) &\leq h(x_i, \gamma) + W \|t_i\|^2 \\
 \Rightarrow \phi(x_i, t_i, \gamma) &\geq 1 - \frac{W \|t_i\|^2}{F(x_i) - h(x_i, \gamma)} \quad . \quad (3.3)
 \end{aligned}$$

Et, par l'inégalité (2.11), nous obtenons

$$\phi(x_i, t_i, \gamma) \geq 1 - \frac{W F(x_i)^2 \beta^2}{\gamma^2 \varepsilon_i \tau_i} \quad . \quad (3.4)$$

De manière évidente, la partie droite tend vers 0 lorsque  $\gamma \rightarrow \infty$ , ce qui montre que l'on peut toujours choisir  $\gamma_i$  assez grand pour satisfaire (3.1).  $\square$

Théorème 3.2.

*La suite  $F(x_i)$  est convergente et*

$$\lim_{i \rightarrow \infty} F(x_i) - h(x_i, \gamma) = 0 \quad .$$

Preuve :

Par le théorème (3.1), nous pouvons toujours choisir  $\gamma_i$  tel que (3.2) soit satisfaite.

Dès lors,

3.13.

$$F(x_{i+1}) \leq F(x_i) - b (F(x_i) - h(x_i, \gamma_i)) \quad (3.5)$$

Ceci implique que  $F(x_i)$  est décroissante et bornée inférieurement, donc convergente.

Réarrangeant (3.5), nous obtenons

$$F(x_i) - h(x_i, \gamma_i) \leq \frac{1}{b} (F(x_i) - F(x_{i+1}))$$

Ceci prouve le théorème.  $\square$

Théorème 3.3 : Convergence.

*Si une sous-suite  $(x_{\pi i})$  dde  $(x_i)$  converge, son point limite est un point stationnaire de  $F(x)$ .*

Preuve :

(1) La suite  $(\gamma_{\pi i})$  est non-bornée :

Soit  $\hat{x}$  le point limite de  $(x_{\pi i})$  et supposons que  $\hat{x}$  n'est pas un point stationnaire de  $F(x)$ .

Nous choisissons une sous-suite  $(x_{\tau i})$  de  $(x_{\pi i})$  telle que  $(\gamma_{\tau i})$  soit strictement croissante.

Par la continuité de  $h(x, \gamma)$  par rapport à  $x$  et l'inégalité (2.11), il existe un voisinage  $V_{\hat{x}}$  de  $\hat{x}$  tel que, pour tout  $x \in V_{\hat{x}}$ ,

$$F(x) - h(x, \gamma) \geq \xi \varepsilon / \lambda \quad (3.6)$$

où  $\xi$  et  $\varepsilon$  dépendent de  $\hat{x}$  et du voisinage  $V_{\hat{x}}$ .

Cependant, par (3.4) ,

$$\phi(x, t, \gamma) \geq 1 - \frac{W \|t\|^2}{F(x) - h(x, \gamma)} .$$

D'autre part, par l'étape 4 de l'algorithme ci-après, nous pouvons choisir  $(\gamma_{\tau i})$  telle que  $\phi(x_{\tau i}, t_{\tau i}, \gamma_{\tau i}) < 1 - b$  .

L'on conclut donc, par le lemme (2.1) , que

$$F(x_{\tau i}) - h(x_{\tau i}, \gamma_{\tau i}) \leq (W/b) F(x_{\tau i})^2 \beta^2 / \gamma_{\tau i}^2 , \quad (3.7)$$

ce qui, pour  $\gamma_{\tau i}$  assez grand, contredit (3.6) .

(2) La suite  $(\gamma_{\tau i})$  est bornée :

On peut donc extraire une sous-suite  $\gamma_{\tau i}$  convergente vers un réel positif ou nul  $\hat{\gamma}$  .

Cependant, la construction de notre algorithme permet toujours de supposer que

$$\gamma_{\tau i} \geq \eta \quad \eta > 0 .$$

En effet, le cas où  $\gamma_{\tau i} \rightarrow 0$  peut conduire à des problèmes du type de l'algorithme de Gauss-Newton, c'est-à-dire la correction  $t_i$  est non-bornée, ce que nous voulons précisément éviter.

Dès lors, nous pouvons supposer  $\hat{\gamma} \neq 0$  .

Par le lemme (2.1) , il existe une constante  $K > 0$  (par exemple  $K = \beta / (\hat{\gamma} - \eta/2)$ ) et un rang  $i_0$  tels que, pour tout  $i \geq i_0$  ,

$$\|t_{\tau i}\|_{B_2} \leq K F(x_{\tau i}) \leq K F(x_0) .$$

Dès lors, un argument du type du corollaire (3.2) p. 2.20, montre que

$$h(x_{\tau i}, \gamma_{\tau i}) \rightarrow h(\hat{x}, \hat{\gamma})$$

et, comme

$$F(x_{\tau i}) \rightarrow F(\hat{x}),$$

nous obtenons, grâce au théorème (3.2),

$$h(\hat{x}, \hat{\gamma}) = F(\hat{x}) \quad .\square$$

### 3.3. Forme de l'algorithme.

---

#### Etape 0.

---

Initialiser  $x_0$ ,  $\hat{\gamma}_0$ ,  $b$ ,  $\theta$ ,  $\varepsilon$ ,  $\chi$ ,  $\tau$ ,  $i := 0$  ( $\chi > 1$ ,  $0 < b < 1/2$ ,  $0 < \theta < 1$ ) ;  $\tau$  est une constante positive dépendant de la précision numérique de l'ordinateur employé.

#### Etape 1.

---

Poser  $j := 1$ .

Calculer  $f(x_i)$ ,  $A_i$ ,  $C_i$  où

$$A_i = A(x_i) \quad \text{et} \quad C_i = C(x_i).$$

Etape 2.

---

Déterminer le vecteur  $t_i(\gamma_i^j)$  minimisant  $H(x_i, t, \gamma_i^j)$  par l'algorithme du chapitre I.

Etape 3.

---

Cette étape analyse les différents cas possibles :

*Si*

$$F(x_i) - h(x_i, \gamma_i^j) < \text{eps} \quad :$$

STOP, on est à l'optimum;

*autrement,*

*si*

$$\phi(x_i, t_i(\gamma_i^j), \gamma_i^j) \geq b \quad ,$$

aller en 4 ;

*autrement,*

augmenter le paramètre  $\gamma_i^j$ , c'est-à-dire

$$j := j + 1$$

$$\gamma_i^j = \chi * \gamma_i^{j-1}$$

et aller en 2 .



#### Etape 4.

Nous verrons que la convergence de l'algorithme semble plus rapide si la suite  $(\gamma_i)$  est bornée; aussi essaierons-nous de contrôler la taille de  $(\gamma_i)$  en imposant que, si  $\gamma_i$  a dû être augmenté dans le sens où  $\gamma_i > \gamma_{i-1}$ , nous tâcherons de limiter la taille de  $\gamma_i$  en imposant

$$\phi(x_i, t_i(\gamma_i), \gamma_i) \leq 1 - b.$$

Nous démontrerons ci-dessous que cette condition peut toujours être satisfaite. Dès lors, nous avons :

Si

$$j = 1,$$

aller en 5 ;

Si

$$\phi(x_i, t_i(\gamma_i^j), \gamma_i^j) < 1 - b,$$

aller en 5 ;

autrement,

(1) poser

$$\alpha = \gamma_i^{j-1}, \quad \beta = \gamma_i^j ;$$

(2) poser

$$\gamma = \frac{1}{2} (\alpha + \beta)$$

et calculer  $t_i(\gamma)$  et  $h(x_i, \gamma)$ .

si

$$(F(x_i) - h(x_i, \gamma)) < \varepsilon \quad :$$

STOP ;

autrement,

aller en (3) ;

(3) si

$$b \leq \phi(x_i, t_i(\gamma), \gamma) \leq 1 - b \quad ,$$

poser

$$j := j + 1$$

$$\gamma_i^j = \gamma$$

et aller en 5 ;

autrement,

$$\text{si} \quad \phi(x_i, t_i(\gamma), \gamma) \geq b \quad ,$$

poser  $\alpha = \gamma$  et aller en (2) ;

autrement,

poser  $\beta = \gamma$  et aller en (2) .

Cette étape est simplement une méthode de bisection pour satisfaire la condition

$$b \leq \phi(x_i, t_i(\gamma_i^j), \gamma_i^j) \leq 1 - b \quad .$$

Etape 5.

Cette étape construit la nouvelle approximation  $x_{i+1}$  ainsi que le paramètre  $\gamma_{i+1}^1$ .

Dès lors, poser

$$\gamma_i = \gamma_i^j, \quad x_{i+1} = x_i + t_i.$$

Si

$$j > 1,$$

poser

$$\gamma_{i+1}^1 = \gamma_i$$

et aller en 1 ;

autrement,

si

$$\gamma_i > \tau,$$

poser

$$\gamma_{i+1}^1 = \gamma_i * \theta ;$$

de cette manière, nous diminuons  $\gamma_{i+1}$  par rapport à  $\gamma_i$  si les conditions

$$b \leq \phi(x_i, t_i(\gamma_i^1), \gamma_i^1) \quad \text{et} \quad \gamma_i^1 > \tau$$

sont satisfaites.

Remarque.

Il nous faut voir que la condition

$$b \leq \phi(x_i, t_i(\gamma_i), \gamma_i) < 1 - b$$

peut toujours être satisfaite lorsque

$$\phi(x_i, t_i(\gamma_i^1), \gamma_i^1) < b,$$

et que cette condition limite la taille du paramètre  $\gamma_i$ .

En effet, puisque

- il existe  $\gamma = \gamma_i^1$  tel que

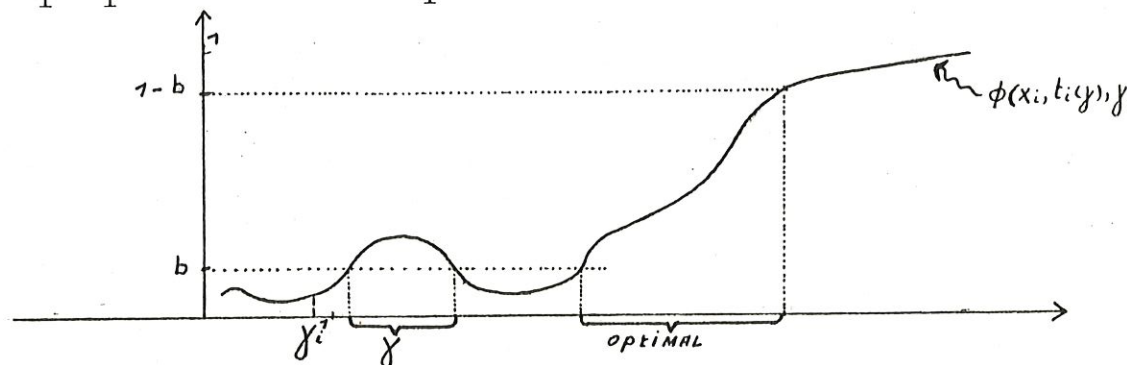
$$\phi(x_i, t_i(\gamma), \gamma) < b,$$

-  $\phi(x_i, t_i(\gamma), \gamma)$  est continu par rapport à  $\gamma$ ,

- par le théorème (3.1),

$$\phi(x_i, t_i(\gamma), \gamma) = 1 + O(1/\gamma^2),$$

la fonction  $\phi(x_i, t_i(\gamma), \gamma)$  a, pour  $x_i$  fixé, la forme suivante



ce qui assure la convergence de la méthode de bisection de l'Etape 4.

#### § 4. VITESSE DE CONVERGENCE DE L'ALGORITHME

---

Nous ne démontrerons pas un théorème de convergence valable dans tous les cas; cependant, voici quelques remarques utiles.

- (1) S'il existe  $i_0$  tel que, pour tout  $i \geq i_0$ ,  $\sigma_i^{(2)}$  est vide et si la règle des multiplicateurs est satisfaite [chap. II, § 5], c'est-à-dire si

$$\left\{ \begin{array}{l} \sigma_i^{(1)} = \sigma_{i_0}^{(1)} \\ \mu_k^{(i)} \geq \mu > 0 \end{array} \right. \quad \forall i \geq i_0 \quad (4.1)$$

$$\left\{ \begin{array}{l} \mu_k^{(i)} \geq \mu > 0 \end{array} \right. \quad \forall i \geq i_0, \quad \forall k \in \sigma_i, \quad (4.2)$$

alors la convergence est quadratique.

Il suffit de remarquer que la correction  $t(\gamma_i)$  de Levenberg est identique à celle de Gauss-Newton dont la convergence quadratique a été prouvée [chap. II, théorème 4.2] .□

- (2) Si la condition (4.1) est satisfaite et si

$$\gamma_i \leq K_2 \|x^* - x_i\|, \quad K_2 > 0, \quad \forall i \geq i_0$$

$$\operatorname{rg} A(x) = p \quad \text{sur } R_c$$

$$F(x_i) \rightarrow 0,$$

alors la convergence est quadratique.

Nous avons

$$f(x^*) = f(x_i) + A_i(x^* - x_i) + \|x^* - x_i\|^2 W_i = 0$$

Or,

$$\begin{aligned} h(x_i, \gamma_i) &\leq \max \{ \|f_i + A_i(x^* - x_i)\|_{B_1}, \gamma_i \|C_i(x^* - x_i)\|_{B_2} \} \\ &\leq \max \{ \|x^* - x_i\|^2 W, \gamma_i \|x^* - x_i\|_{B_2} \} \quad \text{par (2.7)} \\ &\leq K_1 \|x^* - x_i\|^2 \end{aligned} \quad (4.3)$$

où

$$K_1 = \max \{W, K_2\}$$

D'autre part,

$$x_{i+1} = x_i + t_i$$

implique que

$$t_i = (x^* - x_i) - (x^* - x_{i+1})$$

Posons

$$\begin{pmatrix} f_i \\ 0 \end{pmatrix} + \begin{pmatrix} A_i \\ \gamma_i C_i \end{pmatrix} [x^* - x_i - (x^* - x_{i+1})] = \begin{pmatrix} r_i^{(1)} \\ r_i^{(2)} \end{pmatrix}$$

et, par (2.3),

$$f(x_i) + A_i [x^* - x_i - (x^* - x_{i+1})] = -\|x^* - x_i\|^2 W_i, \quad ,$$

d'où

$$A_i(x^* - x_{i+1}) = -r_i^{(1)} - \|x^* - x_i\|^2 W_i \quad (4.4)$$

En inversant l'un des mineurs non-nuls, soit  $Q_i$  de  $A_i^{\sigma_i}$ ,

$$\|x^* - x_{i+1}\| \leq \|Q_i^{-1}\| (\|r_i^{(1)}\| + \|x^* - x_i\|^2 \|W_i\|)$$

Or,

$$\|r_i^{(1)}\| \leq K_3 \|r_i^{(1)}\|_{B_1} \leq h(x_i, \gamma_i) K_3 \leq K_3 K_1 \|x^* - x_i\|^2$$

avec

$$\|W_i\| \leq K_3 \|W_i\|_{B_1} \leq K_3 W$$

Un raisonnement analogue à celui du lemme (3.2) du chapitre précédent permet de trouver une borne uniforme  $K_4$  pour  $\|Q_i^{-1}\|$ .

Dès lors,

$$\|x^* - x_{i+1}\| \leq K_4 (K_3 K_1 + W K_3) \|x^* - x_i\|^2 \quad (4.5)$$

Remarques.

- (1) Les résultats suggèrent qu'il y a intérêt à diminuer  $\gamma$  autant que possible; c'est la raison pour laquelle  $\gamma$  est diminué autant que possible durant l'algorithme.
- (2) Il est intéressant de noter que, lorsque  $(\gamma_i)$  est non-bornée, la décroissance de l'algorithme de Levenberg dépend du carré du paramètre  $\gamma$  alors que, dans l'algorithme de Gauss-Newton, elle est une fonction linéaire de ce paramètre. Ceci résulte d'une comparaison des formules



(3.7) du chapitre III et (3.16) du chapitre II .

Ceci peut entraîner une convergence plus rapide de l'algorithme de Levenberg si  $\|t_i\| \rightarrow 0$  .

## REFERENCES

- 
- [1] D.H. ANDERSON, "Linear Programming and the Calculation of Maximum Norm Approximation", Thèse de Doctorat, Australian National University, avril 1977.
  - [2] J. STOER et C. WITZGALL, "Convexity and Optimization in Finite Dimension", vol. I, Springer, Berlin - Heidelberg - New York, 1970.
  - [3] K. MADSEN, "An Algorithm for Minimax Solution of Overdetermined Systems of Non-Linear Equations", J.J.M.A. 16, 1975, pp. 321-328.

## CHAPITRE IV :

UN ALGORITHME DE LEVENBERG POUR  
LA RÉOLUTION D'UN PROBLÈME DE  
MINIMISATION EN NORME EUCLIDIENNE

## § 1. INTRODUCTION

---

Les algorithmes décrits dans les sections précédentes permettent de résoudre des problèmes de minimisation pour une classe de normes dites polyédrales et comprenant les normes  $l_1$  et  $l_\infty$ . Ils ne s'appliquent malheureusement pas à la troisième norme très usitée, surtout en statistique, la norme  $l_2$ .

Les algorithmes visant à résoudre le problème

$$\min \|f(x)\|_2 \quad (1.1)$$

sont nombreux.

Ceux utilisant une correction de Gauss-Newton cherchent le vecteur correction au point  $x$  en résolvant le problème linéaire

$$\min_{t \in \mathbb{R}^p} \|f(x) + A(x) t\|_2 \quad (1.2)$$

où  $A(x)$  est la matrice jacobienne au point  $x$ .

Malheureusement, cette correction peut conduire à des problèmes concernant la convergence de l'algorithme lorsque  $t(x)$ , vecteur-solution de (1.2), est non-borné, c'est-à-dire lorsque la matrice  $A(x)$  n'est pas de plein rang.

Les algorithmes de Levenberg permettent, comme déjà vu au chapitre III, de résoudre ce problème en considérant, à chaque itération, le problème

$$\min_{t \in \mathbb{R}^p} \|r(t)\|_2$$

où

$$r(t) = \begin{pmatrix} f(x) \\ 0 \end{pmatrix} + \begin{pmatrix} A(x) \\ C(x) \end{pmatrix} t, \quad ,$$

$x$  est fixé et  $C(x)$  une matrice carrée de rang maximal que l'on essaiera de choisir la mieux conditionnée possible (orthogonale, par exemple).

## § 2. QUELQUES RAPPELS SUR LES MOINDRES CARRES

---

Avant d'aborder le problème non-linéaire, rappelons que

Théorème 2.1.

*Une solution du problème*

$$\min_{x \in \mathbb{R}^p} \|A x - b\|_2$$

où  $A(n, p)$ , est donnée par  $x^*$ , solution des équations normales

$$A^T A x = A^T b$$

Preuve :

Posant

$$r = A X - b,$$

nous avons

$$\|r\|_2^2 = r^T r = (-b^t + X^t A^t) (A X - b).$$

D'où,

$$\nabla_X \|r(x)\|_2^2 = A^t A x - A^t b$$

et  $x^*$  est un point minimum

équivalent à

$$\nabla_X \|r(x)\|_2^2 = 0_n \Leftrightarrow A^t A X - A^t b = 0_n$$

□

Si, d'un point de vue théorique, les équations normales sont utiles, elles sont numériquement très mauvaises. Aussi résolvons-nous le problème (2.1) par une méthode inspirée des factorisations de Householder et en factorisant la matrice  $A$  sous la forme de produit de deux matrices :

$$\begin{matrix} A & = & W & \begin{pmatrix} S \\ 0 \end{pmatrix} \\ (n, p) & & (n, n) & (n, p) \end{matrix}$$

où  $W$  est orthogonale et  $S(p, p)$  triangulaire supérieure.

Les détails concernant cette méthode se trouvent, par exemple, dans [1, pp. 59-61] .

Lemme 2.2.

Si  $x^*$  est solution de (2.1) et si  $r(x^*)$  est le vecteur des résidus :

$$r(x^*) = A x^* - b ,$$

alors

$$A^t r(x^*) = 0 .$$

Preuve :

Par définition des équations normales,

$$A^t r(x^*) = A^t A x^* - A^t b = 0 \quad .\square$$



### § 3. ALGORITHME THEORIQUE

---

#### 3.1. Idée et Hypothèse de Travail.

---

Comme vu dans l'introduction, une itération  $i$  aura la forme suivante

(1) Résoudre le problème

$$\left. \begin{array}{l} \min_{h \in \mathbb{R}^p} \|r(h)\|_2 \\ \text{où} \\ r(h) = \begin{pmatrix} f(x_i) \\ 0 \end{pmatrix} + \begin{pmatrix} A(x_i) \\ \gamma C(x_i) \end{pmatrix} h \end{array} \right\} \quad (3.1)$$

où le paramètre  $\gamma$  est choisi afin d'assurer la descente de l'algorithme.

(2) Poser

$$x_{i+1} = x_i + h_i$$

où  $h_i$  est solution de (3.1) .

L'originalité de notre algorithme réside dans le choix du paramètre  $\gamma$ , choix numériquement simple à effectuer et assurant une convergence globale moyennant très peu d'hypothèses.

Les remarques ci-après conduisent aux hypothèses de travail suivantes :

4.6.

$$\phi : \mathbb{R}^p \rightarrow \mathbb{R}^n ; \quad n > p \quad \phi \in C^1 \quad (3.2)$$

$$A(x) = (\nabla \phi_1(x), \dots, \nabla \phi_n(x))^T$$

De plus, à chaque  $x$  de  $\mathbb{R}^p$ , nous associerons une matrice  $C(x)$  de dimensions  $(p, p)$ , non singulière et telle que, pour tout  $x \in \mathbb{R}^p$ ,

$$(1) \quad \|C(x)\|_2 = 1 \quad (3.3)$$

$$(2) \quad \|C^{-1}(x)\|_2 \leq \alpha \quad \alpha > 0$$

$$(3) \quad C(x) \text{ est continue par rapport à } x$$

Dans notre cas, nous choisirons, pour tout  $x \in \mathbb{R}^p$ ,

$$C(x) = \text{id } (p, p)$$

L'algorithme que nous allons envisager construit une suite de points  $(x_i)$  telle que

$$\|f(x_i)\|_2 - \|r(x_i)\|_2 \xrightarrow{i \rightarrow \infty} 0$$

Ceci suggère une caractérisation d'un point minimum un peu différente de celle effectuée sur le vecteur des gradients, à savoir

$$\nabla_X \|f(x)\|_2^2 = 0$$

Lemme 3.1.

Si  $\gamma$  est fini, les assertions suivantes sont équivalentes :

$$(1) \quad \begin{pmatrix} f(x) \\ 0 \end{pmatrix} = h'(x) \quad ;$$

$$(2) \quad \|h'(x)\|_2^2 = \|f(x)\|_2^2 \quad ;$$

(3)  $x$  est un point stationnaire de

$$F(x) = \|f(x)\|_2^2$$

Preuve :

Désignons  $f(x)$  par  $f$  ,  $r'(x)$  par  $r'$  ,  $A(x)$  par  $A$  et  $C(x)$  par  $C$  .

Par le théorème (2.1) , nous avons

$$[A^t A + \gamma^2 C^T C] h = -A^t f \quad (3.4)$$

Dès lors, en posant

$$\begin{pmatrix} r'_1 \\ r'_2 \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} + \begin{pmatrix} A \\ \gamma C \end{pmatrix} h \quad , \quad (3.5)$$

nous obtenons, par le lemme (2.1) ,

$$[A^t, \gamma C^T] r' = 0$$

D'où,

$$f^t r'_1 = \|r'\|_2^2 \quad (3.6)$$

1) (2)  $\Rightarrow$  (1) :

Par Cauchy-Schwartz, on tire de (3.6)

$$\|r'\|_2^2 \leq \|r'_1\|_2 \|f\|_2 = \|r'\|_2 \|r'_1\|_2 \Rightarrow \|r'\|_2 \leq \|r'_1\|_2$$

Comme on a toujours

$$\|r'_1\|_2 \leq \|r'\|_2$$

on obtient

$$\|r'\|_2 = \|r'_1\|_2 \Rightarrow r'_2 = 0$$

Dès lors,

$$f^t r'_1 = \|r'_1\|_2^2 = \|f^t\|_2 \|r'_1\|_2$$

ssi

$$f = r'_1$$

2) (1)  $\Rightarrow$  (3) :

De (1), on tire

$$0 = [f^t, 0] \begin{pmatrix} A \\ \gamma C \end{pmatrix}$$

Or,

$$\nabla \|f(x)\|_2^2 = \nabla (f^T(x) f(x)) = 2 f^T(x) A(x)$$

3) (3)  $\Rightarrow$  (2) :

Multiplions (3.5) par  $[f^t, 0]$  ; nous obtenons alors, par (3.6) ,

$$\|f\|_2^2 + \frac{1}{2} (\nabla_X \|f(x)\|_2^2)^t h = r_1^t f = \|r'\|_2^2$$

Comme  $x$  est point stationnaire de  $f(x)$

$$\nabla \|f(x)\|_2^2 = 0$$

c'est-à-dire

$$\|f\|_2^2 = \|r'\|_2^2$$

□

Soit  $h(x, \gamma)$  solution de

$$\min_{h \in \mathbb{R}^p} \|r(h, x, \gamma)\|$$

où

$$r(h, x, \gamma) = \begin{pmatrix} f(x) \\ 0 \end{pmatrix} + \begin{pmatrix} A(x) \\ \gamma C(x) \end{pmatrix} h \quad (3.7)$$

Posons

$$r(x, \gamma) = r(h(x, \gamma), x, \gamma)$$

Nous avons alors le lemme suivant ;

Lemme 3.2.

Pour tout  $(x, \gamma) \in \mathbb{R}^p \times \mathbb{R}_0^+$  ,  $\|r(x, \gamma)\|_2^2$  est continue en  $(x, \gamma)$  .

Preuve :

Rappelons une formule bien connue dont la démonstration se trouve dans [2] .

Formule de Scherman-Morrisson :

Soient  $A(n, n)$  inversible et  $U(n, m)$  et  $V(n, m)$  des matrices.

Si  $(A + U V^T)$  est inversible, nous avons

$$(A + U V^T)^{-1} = A^{-1} - A^{-1} U (I + V^T A^{-1} U)^{-1} V^T A^{-1} \quad (3.8)$$

Dès lors,  $h(x, \gamma)$ , solution de (3.7), c'est-à-dire de (3.4), est donnée par

$$h(x, \gamma) = -[C^T(x) C(x) / \gamma^2 - [I + (A(x) / \gamma^2) [C^T(x) C(x)]^{-1} A(x)]^{-1}] A^t(x) f(x) .$$

Ceci prouve la continuité de  $h(x, \gamma)$  par rapport à  $(x, \gamma)$  pour tout  $(x, \gamma) \in \mathbb{R}^p \times \mathbb{R}_0^+$  puisque  $A(x)$ ,  $C(x)$  et  $f(x)$  sont continues par rapport à  $x$ .

De ceci, on déduit immédiatement la continuité de  $r(x, \gamma)$  et donc de

$$\|r(x, \gamma)\|_2^2 = r^t(x, \gamma) \cdot r(x, \gamma) \quad \square$$

Afin de construire notre algorithme, nous introduisons la fonction

$$\psi(x_i, \gamma) = \frac{F(x_i) - F(x_{i+1})}{2 [F(x_i) - \|r(x_i, \gamma_i)\|_2^2]} \quad (3.9)$$

où

$$F(x) = \|f(x)\|_2^2, \quad ,$$

$$x_{i+1} = x_i + h_i(x_i, \gamma_i) \quad ,$$

$h(x_i, \gamma_i)$  étant le vecteur solution de (3.7) .

Notre but sera de déterminer, en chaque point  $x_i$  , un paramètre  $\gamma_i$  tel que

$$\psi(x_i, \gamma_i) \geq \sigma \quad (3.10)$$

où  $\sigma$  est une constante positive indépendante de  $x_i$  .

Si nous réécrivons la condition (3.10), nous obtenons

$$F(x_{i+1}) = F(x_i) - 2 \sigma (F(x_i) - \|r(x_i, \gamma_i)\|_2^2) \quad ,$$

ce qui, puisque l'on a toujours

$$F(x_i) \geq \|r(x_i, \gamma_i)\|_2^2$$

si  $x_i$  n'est pas optimal, exprime que notre algorithme sera un algorithme de descente.

Naturellement, il nous faut vérifier que cette condition peut toujours être vérifiée, ce sera là le but du théorème suivant qui nous fournira, par ailleurs, le moyen de déterminer le paramètre  $\gamma_i$  .



### 3.2. Convergence.

---

#### Théorème 3.2.

---

Pour tout  $\sigma < 1$  et pour tout  $x \in \mathbb{R}^p$ , il existe  $\gamma$  tel que la relation (3.10) soit satisfaite.

Preuve :

Nous allons prouver que  $\psi(x, \gamma) \rightarrow 1 + O(1/\gamma^2)$ , si bien que le test (3.10) peut être satisfait en prenant  $\gamma$  assez grand.

Réappliquant la formule de Sherman-Morrison, nous obtenons

$$h(x_i, \gamma_i) = -[C^T(x_i) C(x_i) / \gamma_i^2 - I + (A(x_i) / \gamma_i^2) [C^T(x_i) C(x_i)]^{-1} A(x_i)]^{-1} A^t(x_i) f(x_i) .$$

D'où,

$$h(x_i, \gamma_i) = O(1/\gamma_i^2) \quad (3.11)$$

A présent,

$$\begin{aligned} (\nabla_X F(x_i))^T h(x_i, \gamma_i) &= 2 f^t(x_i) A(x_i) h(x_i, \gamma_i) \\ &= -2 f^t(x_i) [f(x_i) - r_1(x_i)] \end{aligned} \quad (3.12)$$

où

$$r_1(x_i) = [(r(x_i))_1, \dots, (r(x_i))_p]$$

et

4.13.

$$r(x_i) = r(x_i, \gamma_i) \quad (3.12)$$

Or, par le lemme (2.2) ,

$$[f^t(x_i), 0] r(x_i) = r^T(x_i) r(x_i) \quad (3.13)$$

(3.12) s'écrit donc

$$(\nabla_X F(x_i))^T h(x_i, \gamma_i) = -2 (\|f^t(x_i)\|_2^2 - \|r(x_i)\|_2^2) \quad (3.14)$$

Un développement de  $F(x)$  au premier ordre donne

$$\frac{F(x_i) - F(x_{i+1})}{2 \|F(x_i) - r_i\|_2^2} = - \frac{1}{2 (F(x_i) - \|r_i\|_2^2)} \{2 f_i^t A_i h_i(\gamma) + O(\|h(x_i, \gamma_i)\|_2^2)\}$$

$$= 1 + \frac{O(\|h(x_i, \gamma_i)\|_2^2)}{(\nabla_X F(x_i))^T h(x_i, \gamma_i)} \quad \text{par (3.14)}$$

$$= 1 + O(\|h(x_i, \gamma_i)\|_2)$$

$$= 1 + O(1/\gamma_i^2) \quad \text{par (3.11)} \quad \square$$

Remarquons encore que, grâce à ce théorème, il n'y a aucune restriction à supposer  $\gamma_i \geq \eta > 0$  ( $\eta$  constante fixée indépendante de  $i$ ). Cette restriction est semblable à celle effectuée au chapitre III, p. 3.14 et permet d'assurer que l'algorithme converge vers un point stationnaire, ce qui sera le but du théorème suivant.

Dans ce qui suit, nous poserons

$$r(x_i, \gamma_i) = r_i$$

Théorème 3.1.

Si (1) la suite  $(\gamma_i)$  est bornée inférieurement et supérieurement,

(2) pour tout  $i$ , on a

$$0 < \sigma \leq \psi(x_i, \gamma_i) \quad (3.15)$$

où  $\sigma$  est une constante indépendante de  $i$ ,

alors la suite  $\{F(x_i)\}$  est convergente et le point limite de toute sous-suite  $(x_{\pi i})$  de  $(x_i)$  est un point stationnaire de  $F(x)$ .

Preuve :

Remarquons que la condition (2) et le caractère borné inférieurement de la suite  $(\gamma_i)$  peuvent toujours être vérifiés par le théorème (3.2) et la remarque l'accompagnant. Par conséquent, la seule hypothèse non-assurée est le caractère borné supérieurement de  $(\gamma_i)$ .

Montrons donc que, sous cette hypothèse, le théorème est vérifié et pour cela, notons que, si  $x_i$  n'est pas un point stationnaire de  $f(x)$ , par (3.1), nous aurons

$$\|f(x_i)\|_2 > \|r_i\|_2$$

De plus, si (3.15) est vérifiée, nous avons

$$F(x_{i+1}) < F(x_i)$$

De (3.15), nous tirons aussi

$$\sigma \leq \frac{(\|f(x_i)\|_2 + \|f(x_{i+1})\|_2)(\|f(x_i)\|_2 - \|f(x_{i+1})\|_2)}{2(\|f(x_i)\|_2 + \|r_i\|_2)(\|f(x_i)\|_2 - \|r_i\|_2)}$$

Ceci implique que

$$\sigma \leq \frac{\|f(x_i)\|_2 - \|f(x_{i+1})\|_2}{\|f(x_i)\|_2 - \|r_i\|_2} \quad (3.16)$$

car

$$\|f(x_i)\|_2 \leq \|f(x_i)\|_2 + \|r_i\|_2$$

et

$$\|f(x_i)\|_2 + \|f(x_{i+1})\|_2 \leq 2\|f(x_i)\|_2$$

Et (3.16) implique que

$$\|f(x_{i+1})\|_2 \leq \|f(x_i)\|_2 - \sigma(\|f(x_i)\|_2 - \|r_i\|_2)$$

Ceci implique que la suite  $\|f(x_i)\|_2$  est décroissante et bornée inférieurement, donc convergente.

De plus, (3.16) donne aussi

$$\|f(x_i)\|_2 - \|r_i\|_2 \leq (\|f(x_i)\|_2 - \|f(x_{i+1})\|_2) / \sigma$$

qui implique que

$$\|f(x_i)\|_2 - \|r_i\|_2 \xrightarrow{i \rightarrow \infty} 0$$

A présent, supposons  $(x_{\pi i}) \rightarrow x^*$  où  $(x_{\pi i})$  est une sous-suite de  $(x_i)$ .

Comme  $(\gamma_{\pi i})$  est bornée, on peut extraire de  $(\gamma_{\pi i})$  une sous-suite  $(\gamma_{\tau i})$

convergente vers  $\gamma^* \neq 0$ , car  $(\gamma_{\tau i})$  est bornée inférieurement par une constante  $\eta > 0$ .

Dès lors,  $(x_{\tau i}) \rightarrow x^*$  et  $(\gamma_{\tau i}) \rightarrow \gamma^*$  avec  $\gamma^* > \eta > 0$ .

Par le lemme (3.2) et le résultat ci-dessus, nous avons

$$\|f(x_{\tau i})\|_2 - \|r_{\tau i}\|_2 \rightarrow \|f(x^*)\|_2 - \|r^*\|_2$$

et donc

$$\|f(x^*)\|_2 = \|r(x^*, \hat{\gamma})\|_2,$$

ce qui achève la démonstration compte-tenu du Lemme 3.1.  $\square$

### 3.3. Algorithme.

---

N.B. : Nous prendrons, pour tout  $x \in \mathbb{R}^p$ ,

$$C(x) = I_p.$$

Etape 0 : Initialisations.

---

Initialiser  $x_0$ ,  $\gamma_0^1$ ,  $\sigma$ ,  $\text{expn}$ ,  $\text{decr}$ ,  $\text{eps}$ ,  $\text{gmax}$ ,  $\theta$  avec

$$0 < \sigma < 1, \quad 0 < \text{decr} < 1$$

$$\text{expn} > 1,$$

$\theta$  borne inférieure pour  $\gamma$ , pouvant dépendre de la précision numérique de l'ordinateur, et  $i := 0$ .

Etape 1.

---

Poser

$$j := 1$$

et calculer  $f(x_i)$  et  $A(x_i)$ .

Etape 2.

---

Factoriser  $A(x_i)$  telle que

$$\begin{matrix} A(x_i) & = & W_i & \begin{pmatrix} S_i \\ 0 \end{pmatrix} \\ (n, p) & & (n, n) & (n, p) \end{matrix}$$

où  $W_i(n, n)$  est orthogonale et  $S_i(p, p)$  est triangulaire supérieure.

Ceci est toujours possible [cf. [1], pp. 59-60].

Calculer

$$f_1(x_i) = W_i^T f(x_i)$$

Etape 3.

---

Factoriser  $\begin{pmatrix} S_i \\ 0 \\ \gamma_i^j \text{ Id} \end{pmatrix}$  telle que

$$\begin{pmatrix} S_i \\ 0 \\ \gamma_i^j \text{ Id} \end{pmatrix}_{(n+p, p)} = Q_i \begin{pmatrix} U_i \\ 0 \end{pmatrix}_{(n+p, n+p)} \quad (n+p, p)$$

où  $Q_i(n+p, n+p)$  est orthogonale et  $U_i(p, p)$  est triangulaire supérieure.

Calculer

$$\begin{pmatrix} f_2(x_i) \\ (p, 1) \\ f_3(x_i) \\ (n, 1) \end{pmatrix} = Q_i^T \begin{pmatrix} f_1(x_i) \\ (n, 1) \\ 0 \\ (p, 1) \end{pmatrix}$$

Etape 4.

---

Calculer

$$h(x_i, \gamma_i^j) = -U_i^{-1} f_2(x_i)$$

et

$$\|r(x_i, \gamma_i^j)\|_2^2 = \|f_3(x_i)\|_2^2$$

Etape 5.

---

Calculer  $F(x_i + h(x_i, \gamma_i^j))$  et  $\psi(x_i, \gamma_i^j)$ .



*Si*

$$F(x_i) - \|r(x_i, \gamma_i^j)\|_2^2 \leq \text{eps} \quad :$$

STOP : on est à l'optimum;

*autrement,*

*si*

$$\psi(x_i, \gamma_i^j) \geq \sigma \quad ,$$

aller en 6.;

*autrement,*

augmenter  $\gamma_i^j$  :

$$j := j + 1$$

$$\gamma_i^j = \text{expn} * \gamma_i^{j-1} \quad ;$$

ce test permet de déterminer si le paramètre  $\gamma$  n'est pas trop grand :

*si*

$$\gamma_i^j > g_{\max} \quad :$$

STOP : non convergence possible;

*autrement,*

aller en 3.

Etape 6.

Construire la nouvelle approximation  $x_{i+1}$  et déterminer le paramètre  $\gamma_{i+1}^1$ .

$$x_{i+1} = x_i + h(x_i, \gamma_i^j)$$

Si

$$j = 1 \quad \text{et} \quad \gamma_i^j > \theta,$$

alors

$$\gamma_{i+1}^1 = \gamma_i^j * \text{decr} ;$$

autrement,

$$\gamma_{i+1}^1 = \gamma_i^1,$$

poser

$$i := i + 1$$

et aller en 1.

A l'Etape 6, nous essayons de diminuer  $\gamma$  autant que possible; ceci peut se justifier puisqu'il est connu que le cas  $\gamma = 0$  correspondant à l'itération de Gauss-Newton conduit à un vecteur-correction  $h(x_i, 0)$  assez bon.

Le paragraphe suivant soulignera également ce point.

#### § 4. REMARQUES COMPLEMENTAIRES

---

Nous allons démontrer, dans un cas particulier, que la vitesse de convergence est linéaire et même quadratique si  $\gamma_i$  peut être choisi assez petit.

##### Théorème 4.1.

S'il existe  $j$  tel que

(1) l'ensemble  $R = \{x \mid F(x) \leq F(x_j)\}$  soit borné;

(2) pour tout  $x, y \in R$ ,

$$f(y) = f(x) + (y - x) + \|y - x\|_2^2 w(x, y)$$

où  $w(x, y)$  est un vecteur dépendant de  $(x, y)$  tel que

$$\|w(x, y)\|_2^2 \leq w; \quad ;$$

(3)  $A(x)$  est de rang maximum dans  $R$ ;

et si

$$f(x^*) = \theta_n \quad ;$$

alors, s'il existe  $K > 0$  tel que

$$\gamma_i \leq K \|x^* - x_i\|_2^{1/2},$$

la convergence est linéaire;

s'il existe  $K' > 0$  tel que

$$\gamma_i \leq K' \|x^* - x_i\|_2,$$

la convergence est quadratique.

Preuve :

Dans tout ce qui suit, nous supposons  $i \geq j$ .

Puisque  $x_i \rightarrow x^*$ , le paramètre  $\gamma_i$  est borné supérieurement par une constante  $M$ .

Posons

$$K_1 = \sup_{\substack{x \in R \\ 0 \leq \gamma \leq M}} \| (A^t(x) A(x) + \gamma^2 \text{Id})^{-1} \|_2$$

$$K_2 = \sup_{x \in R} \|A^t(x)\|_2$$

$$K_3 = \sup_{x \in R} \| (A^t(x) A(x))^{-1} \|_2$$

Dès lors,

$$\|h(x_i, \gamma_i)\|_2 \leq K_1 K_2 \|f(x_i)\|_2$$

De plus,

$$\theta_n = f(x^*) = f(x_i) + A_i (x^* - x_i) + \|x^* - x_i\|_2^2 W(x_i, x^*)$$

qui implique

$$\theta_p = A_i^t f(x_i) + A_i^t A_i (x^* - x_i) + A_i^t \|x^* - x_i\|_2^2 W(x_i, x^*)$$

Or,

$$A_i^t f(x_i) + (A_i^t A_i + \gamma_i^2 \text{Id}) h(x_i, \gamma_i) = \theta_p$$

D'où

$$\theta_p = A_i^t A_i (x^* - (x_i + h(x_i, \gamma_i))) + \gamma_i^2 h(x_i, \gamma_i) + A_i^t \|x^* - x_i\|_2^2 W(x_i, x^*),$$

c'est-à-dire

$$-(A_i^t A_i) (x^* - x_{i+1}) = \gamma_i^2 h(x_i, \gamma_i) + A_i^t \|x^* - x_i\|_2^2 W(x_i, x^*) \quad (4.1)$$

Donc, si

$$\gamma_i \leq K \|x^* - x_i\|_2^{1/2},$$

on a

$$\|x^* - x_{i+1}\|_2 \leq (K_1 K_2 K_3 K^2 \|f(x_j)\|_2) \|x^* - x_i\|_2 + K_2 K_3 W \|x^* - x_i\|_2^2,$$

ce qui démontre la convergence linéaire.

Et, si

$$\gamma_i \leq K' \|x^* - x_i\|_2,$$

on obtient, à partir de (4.1),

$$\|x^* - x_{i+1}\|_2 \leq (K_1 K_2 K_3 K'_2 \|f(x_j)\|_2 + K_2 K_3 W) \|x^* - x_i\|_2^2,$$

ce qui démontre la convergence quadratique.  $\square$

Pour terminer, analysons plus en détail le cas où la suite  $(\gamma_i)$  est non-bornée, c'est-à-dire lorsque la convergence de l'algorithme vers un point stationnaire est non-assurée. Nous allons voir que cette hypothèse est fortement liée au

Hessien de  $\|f(x)\|_2^2$ , à savoir

Théorème 4.2.

Si  $(\gamma_i)$  est non-bornée et si  $f \in C^2(\mathbb{R}^p, \mathbb{R}^n)$ , alors  $\|\nabla^2 F(x)\|_2$  est non-bornée.

Preuve :

Si  $(\gamma_i)$  est non-bornée, par construction de notre algorithme, il existe une sous-suite  $(\gamma_i^*)$  non-bornée telle que

$$\sigma > \psi(x_i, \gamma_i^* / \exp n) \quad (4.2)$$

Posons

$$\gamma_i^* / \exp n = \tilde{\gamma}_i$$

Nous avons donc

$$\sigma > \psi(x_i, \tilde{\gamma}_i)$$

Comme  $F(X)$  est de classe  $C^2$ , nous obtenons

$$F(x_i) - F(\tilde{x}_{i+1}) = -\nabla F_i h(x_i, \tilde{\gamma}_i) - \frac{1}{2} h^t(x_i, \tilde{\gamma}_i) \overline{\nabla^2} F(x_i) h(x_i, \tilde{\gamma}_i)$$

où

$$\tilde{x}_{i+1} = x_i + h(x_i, \tilde{\gamma}_i)$$

et où  $\overline{\nabla^2} F(x_i)$  désigne la valeur moyenne du Hessien.

Or, pour tout  $a, b \in \mathbb{R}$ ,

$$|a + b| \geq |b| - |a|$$

Dès lors, (4.2) s'écrit

$$\begin{aligned} \sigma &> \frac{|\nabla F_i h(x_i, \tilde{\gamma}_i)| - \frac{1}{2} h^t(x_i, \tilde{\gamma}_i) \sqrt{\nabla^2 F_i h(x_i, \tilde{\gamma}_i)}}{|\nabla F_i h(x_i, \tilde{\gamma}_i)|} \\ \Rightarrow |h^t(x_i, \tilde{\gamma}_i) \sqrt{\nabla^2 F_i h(x_i, \tilde{\gamma}_i)}| &> 2(1 - \sigma) |\nabla F_i h(x_i, \tilde{\gamma}_i)| \\ \Rightarrow \|\sqrt{\nabla^2 F_i}\|_2 &> 2(1 - \sigma) \frac{|\nabla F_i h(x_i, \tilde{\gamma}_i)|}{\|h(x_i, \tilde{\gamma}_i)\|_2^2} \end{aligned} \quad (4.3)$$

D'autre part,

$$\begin{aligned} \nabla F_i h(x_i, \tilde{\gamma}_i) &= 2 f_i^t A_i h(x_i, \tilde{\gamma}_i) \\ &= -2 h(x_i, \tilde{\gamma}_i) [A_i^t A_i + \tilde{\gamma}_i^2 C_i^t C_i] h(x_i, \tilde{\gamma}_i) \end{aligned}$$

D'après (3.3), nous concluons donc

$$|\nabla F_i h(x_i, \tilde{\gamma}_i)| \geq 2 \tilde{\gamma}_i^2 \|h(x_i, \tilde{\gamma}_i)\|_2^2$$

En utilisant (4.3), nous obtenons

$$\|\sqrt{\nabla^2 F_i}\|_2 > 4(1 - \sigma) \tilde{\gamma}_i^2,$$

ce qui démontre la thèse puisque la suite  $(\gamma_i)$  est non-bornée.  $\square$



## CONCLUSIONS

---

Notre algorithme diffère des autres algorithmes de type Levenberg par le choix du paramètre  $\gamma$  déterminant la correction au point actuel.

Notre choix de  $\gamma$  possède deux avantages :

- (1) il permet une convergence globale de l'algorithme moyennant une hypothèse très souvent remplie expérimentalement;
- (2) il est aisé à effectuer numériquement.

Remarquons que les exemples traités montrent que l'algorithme est compétitif au point de vue temps calcul par rapport à un algorithme sans utilisation des dérivées de Harwell.

Enfin, la question peut se poser de déterminer les valeurs des paramètres  $\text{expn}$  et  $\text{decr}$  assurant la meilleure convergence; ceci sera discuté dans le chapitre "Résultats Numériques".

REFERENCES

---

- [1] J. KOWALIK et M.R. OSBORNE, "Methods for Unconstrained Optimization Problems", Elsevier, New-York, 1973.
- [2] ORTEGA et RHEINBOLDT, "Iterative Solution of Nonlinear Equations in Several Variables", Academic Press; New York, 1970.
- [3] M.R. OSBORNE, "Some Aspects of Nonlinear Least Squares Calculations", dans "Numerical Methods for Non-Linear Optimization", ed. F.A. Lootsma, Academic Press, 1972.
- [4] E.M.L. BEALE, "Numerical Methods", dans "Nonlinear Programming", ed. J. Abadie, North Holland, 1967.
- [5] D.W. MARQUADT, "An Algorithm for Least Squares Estimation of Non-Linear Parameters", Siam J. Appl. Math., 11, 1963, pp. 431-441.
- [6] W.C. DAVIDON, "New Least-Squares Algorithms", dans "J. Optimization Theory Appl.", 18, 1976, pp. 187-197.
- [7] G.H. GOLUB, "Matrix Decompositions and Statistical Calculations", dans "Statistical Computation", ed. R.C. Milton et J.A. Nelder, Academic Press, New York, 1969.
- [8] J.E. DENNIS et J.J. MORE, "Quasi-Newton Methods - Motivation and Theory", Siam Revue, 19, 1977, pp. 46-89.

## CHAPITRE V :

### IMPLÉMENTATION ET RÉSULTATS NUMÉRIQUES

## § 1. NORMES POLYEDRALES

---

Les algorithmes relatifs aux normes polyédrales sont les suivants :

- (1) une sous-routine POL, comprenant l'algorithme décrit au chapitre I. Celui-ci résout le problème

$$\min_{h \in \mathbb{R}^p} \|f + A h\|_B$$

où B est une matrice définissant une norme polyédrale; dans notre cas la norme  $l_1$  ou  $l_\infty$ ,

f est un vecteur de dimension n ( $n \leq 50$ ), donné,

A est une matrice de dimension  $(n \times p)$  ( $n > p$ ,  $p \leq 25$ ), donné.

- (2) Une sous-routine MINPOL, comprenant l'algorithme décrit au chapitre II. Celui-ci résout donc le problème

$$\min_{x \in \mathbb{R}^p} \|f(x)\|_B$$

où B est une matrice définissant une norme polyédrale; dans notre cas la norme  $l_1$  ou  $l_\infty$ ,

$f(x)$  est une fonction dérivable de  $\mathbb{R}^p \rightarrow \mathbb{R}^n$  ( $n > p$ ,  $n \leq 50$ ,  $p \leq 25$ ).

La fonction  $f(x)$  et l'approximation  $x_0$  sont supposées vérifier les hypothèses du chapitre II.

Enfin, ces sous-routines sont complétées par une sous-routine AUX fournissant les quantités nécessaires au démarrage de la sous-routine POL, à savoir

(1) une référence initiale  $Z$  de dimensions  $(p+1, n)$  telle que

$$\text{rg}(Z A) = p+1,$$

(2) un vecteur  $v$  des multiplicateurs, positifs ou nuls, tel que

$$v^T (Z A) = (0 \quad \dots \quad 0)$$

$$v^T v = 1$$

### 1.1. Problèmes Linéaires.

---

La sous-routine POL a été utilisée pour résoudre deux problèmes

(1-1.1) Trouver la meilleure approximation au sens de la norme  $l_1$  et  $l_\infty$  de la fonction

$$\exp(\sin(t))$$

par un polynôme de la forme

$$x_1 + x_2 t + x_3 t^2 + x_4 t^3 + x_5 t^4 + x_6 t^5$$

sur l'intervalle discret formé des points

$$-1, -0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8, 1$$

Les résultats numériques, relatifs à ce problème sont :



TABLEAUX I : Approximation de  $\exp(\sin(t))$  par un polynôme du sixième degré sur un ensemble de dix points situés dans l'intervalle  $[-1, 1]$

Norme  $l_1$  :

Itération	X						Valeur de $\lambda^{(i)}$
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	
1	-0.962326	-2.000976	-1.876789	0.180089	1.394914	0.738624	0.001634
2	-1.041559	-2.006217	-1.588317	0.721085	1.199188	0.181238	0.007837
3	-1.022049	-2.144888	-1.659946	0.914643	1.268058	0.195470	0.017648
4	-1.059500	-2.023118	-1.406998	0.380059	1.009891	0.613406	0.043243
5	-1.134917	-2.397602	-0.824126	0.734716	0.440047	-1.267661	0.012076
6	-1.022440	-2.026830	-1.603054	0.628423	1.195464	0.380680	0.075872
7	-0.993603	-1.93177	-1.802755	0.088416	1.389137	0.803280	0.079631
8	-1.012082	-2.025836	-1.643048	0.577130	1.230499	0.401894	0.095190
9	-1.013237	-2.048047	-1.747205	0.511212	1.316852	0.491735	0.101494
10	-1.037389	-2.049001	-1.656099	0.498563	1.260055	0.510353	0.105166
11	-1.031066	-2.065050	-1.640357	0.622673	1.233408	0.398400	0.105861
12	-1.047368	-2.116493	-1.589069	0.801699	1.185436	0.270264	0.116692
13	-1.020437	-2.006281	-1.724783	0.354646	1.302527	0.612486	0.116786
14	-1.022684	-2.027305	-1.699842	0.456777	1.277421	0.529484	0.116833
15	-0.999344	-1.962559	-1.782115	0.237612	1.339089	0.684741	0.117057
16	-1.004486	-1.990004	-1.748846	0.357994	1.308446	0.589982	0.117068
17	-1.029851	-2.125379	-1.584736	0.951816	1.157292	0.122558	0.117126
18	-1.010099	-2.011107	-1.738088	0.432934	1.304746	0.536040	0.118409
19	-1.013602	-2.026569	-1.724753	0.494327	1.294729	0.489553	0.118902
20	-1.009204	-2.014956	-1.741892	0.435373	1.308404	0.539697	0.119252

Le vecteur solution est donc

$$\begin{cases} X_1 = -1.009204 \\ X_2 = -2.014956 \\ X_3 = -1.741892 \\ X_4 = 0.435373 \\ X_5 = 1.308404 \\ X_6 = 0.539697 \end{cases}$$

La valeur de la norme est 0.119252 .

Le vecteur des multiplicateurs à l'optimum est 0.531250 0.125  
0.1875 0.03125 0.0625 0.03125 0.03125 .

La référence optimale est

-1	1	1	-1	1	-1	1	1	-1	1
-1	1	-1	-1	-1	-1	1	1	-1	1
-1	1	1	-1	-1	-1	1	-1	-1	-1
-1	1	-1	-1	1	-1	1	1	-1	-1
-1	1	-1	-1	1	1	1	1	-1	-1
1	1	1	-1	1	-1	1	-1	-1	-1
1	1	1	-1	-1	-1	1	1	-1	-1



Norme  $l_\infty$  :

Itération	X						Valeur de $\lambda$ (i)
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	
1	-1.000394	-2.000976	-1.876789	0.180089	1.394849	0.738624	0.000817
2	-1.005297	-2.005700	-1.850407	0.313013	1.579535	0.820287	0.000962
3	-1.014108	-2.006067	-1.616369	0.675540	1.235363	0.246668	0.003605
4	-1.015197	-1.999611	-1.586268	0.636715	1.163383	0.323010	0.004610
5	-1.056976	-2.190775	-1.567965	1.183770	1.193886	-0.032880	0.011638
6	-1.009073	-2.142486	-1.754747	1.068413	1.341439	0.034187	0.020910
7	-1.018515	-2.119406	-1.664338	0.907358	1.263671	0.172162	0.023511



(1-1.2) Trouver la meilleure approximation au sens de la norme  $l_1$  et  $l_\infty$  de la fonction

$$\arctang(\sin(t))$$

par un polynôme de la forme

$$x_1 + x_2 t^2 + x_3 t^4 + x_4 t^6$$

sur l'ensemble discret formé des points

$$-1, -0.6, -0.2, 0, 0.2, 0.6, 1.$$

Cet exemple montrera l'efficacité de l'algorithme lorsque la condition de Haar n'est pas vérifiée.

Rappelons qu'une matrice  $A(n, p)$  ( $n \geq p$ ) vérifie la condition de Haar si et seulement si toute sous-matrice  $A|(p, p)$  de  $A$  est de rang  $p$ .

Il est manifeste que, dans ce cas-ci, la matrice des variables approximantes n'est pas de Haar.

TABLEAUX II : Approximation de  $\arctang(\sin(t))$  par un polynôme de la forme  $X_1 + X_2 t^2 + X_3 t^4 + X_5 t^6$  sur un ensemble de sept points situés dans l'intervalle  $[-1, 1]$

Norme  $l_1$  :

Itération	X				Valeur de $\lambda(i)$
	$X_1$	$X_2$	$X_3$	$X_4$	
1	-5.11893	171.44576	-583.62272	419.02347	1.0280
2	0.	-12.69376	38.47059	-25.32901	1.0280
3	0.	7.40412	-28.62766	20.97183	1.0280
4	0.	-2.09575	89.44977	-67.79269	1.0280
5	-0.59002	13.61384	-41.15337	28.23906	1.6180
6	1.20125	-64.17446	232.12506	-169.04236	1.6180
	0.19611	-5.20882	7.66970	-0.53721	2.6232
8	0.19611	-5.25130	8.84971	-3.48723	2.6232
9	0.19611	-5.45490	14.12713	-8.16882	2.6232
10	0.19611	-5.45595	14.25482	-8.48802	2.6232
11	0.	6.04084	-29.41160	24.07028	2.8193

Le vecteur solution est donc

$$\begin{cases} X_1 = 0. \\ X_2 = 6.04084 \\ X_3 = -29.4116 \\ X_4 = 24.07028 \end{cases}$$

La valeur de la norme est 2.8193 .

La référence optimale est

-1	-1	-1	1	1	1	-1
-1	-1	1	1	1	1	1
-1	-1	-1	1	-1	1	1
1	-1	-1	1	-1	1	1
1	-1	-1	1	1	1	1
-1	-1	-1	1	1	1	1

Norme  $1_{\infty}$  :

Itération	X				Valeur de $\lambda(i)$
	$X_1$	$X_2$	$X_3$	$X_4$	
1	0.69952	0.54503	-13.99944	7.84961	0.69952
2	0.69952	6.00805	-28.50075	21.79317	0.69952
3	0.69952	-5.48310	14.91028	-10.12670	0.69952

Le vecteur solution est donc

$$\begin{cases} X_1 = 0.69952 \\ X_2 = -5.48310 \\ X_3 = 14.91028 \\ X_4 = -10.12670 \end{cases}$$

La valeur de la norme est 0.6995 .

La référence optimale est

-1	0	0	0	0	0	0
0	0	0	0	1	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1
0	0	0	1	0	0	0

## 1.2. Problèmes Non-Linéaires.

---

La sous-routine MINPOL a été utilisée pour résoudre deux problèmes :

(1-2.1) Trouver la meilleure approximation au sens de la norme  $l_1$  et  $l_\infty$  de la fonction

$$\exp(t)$$

par une fonction de la forme

$$\frac{x_1 + x_2 t}{1 + x_3 t + x_4 t^2 + x_5 t^3}$$

sur l'ensemble discret formé de vingt-et-un points équipartis dans l'intervalle  $[-1, 1]$ ,  $-1$  et  $1$  faisant partie de cet ensemble.

Les valeurs des paramètres utilisés pour résoudre ce problème sont

$$\text{TETA} = 1.E - 01$$

$$\text{EPS} = 1.E - 06$$

$$\text{SIGMA} = 1.E - 04$$

;

GAM désigne le paramètre servant à déterminer le vecteur de descente.



TABLEAUX III : Approximation de  $\exp(t)$  par une fonction de la forme  $(x_1 + x_2 t) / (1 + x_3 t + x_4 t^2 + x_5 t^3)$  sur un ensemble de vingt-et-un points équi-partis dans l'intervalle  $[-1, 1]$

Norme  $l_1$  :

Itération	X					Valeur de la norme	Valeur de GAM
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$		
1	0.5	0.5	0.	0.	0.	14.56669	0.1
2	0.55	0.51311	-0.07344	-0.02650	-0.00863	13.04882	0.1
3	0.595	0.52149	-0.13373	-0.03868	-0.00560	11.70545	1.
4	0.99975	0.56973	-0.63591	-0.08617	-0.05356	6.49396	1.
5	1.00155	0.24694	-0.75183	0.29557	-0.13888	0.82398	1.
6	1.00051	0.19441	-0.80605	0.31259	-0.07107	0.04552	1.
7	0.99993	0.25301	-0.74705	0.24618	-0.03843	0.02276	1.
	0.99990	0.25683	-0.74321	0.24193	-0.03639	0.00156	



Norme  $l_{\infty}$  :

Itération	X					Valeur de la norme	Valeur de GAM
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$		
1	0.5	0.5	0.	0.	0.	1.71828	0.1
2	0.55001	0.51316	-0.07313	-0.02664	-0.00882	1.52558	0.1
3	0.59498	0.52151	-0.13376	-0.03897	-0.00532	1.36007	0.1
4	0.63546	0.52626	-0.18420	-0.04347	-0.00065	1.21537	1.
5	0.99987	0.54457	-0.61353	-0.03945	0.05535	1.12006	1.
6	1.00038	0.32217	-0.67486	0.18801	-0.05109	0.14408	1.
7	1.00005	0.24969	-0.75098	0.25178	-0.04090	0.00140	1.
	0.99987	0.25353	-0.74666	0.24525	-0.03751	0.00130	

Le vecteur solution de la norme  $l_1$  est donc

$$\begin{cases} X_1 = 0.99990 \\ X_2 = 0.25683 \\ X_3 = -0.74321 \\ X_4 = 0.24193 \\ X_5 = -0.03639 \end{cases}$$

La valeur de la norme  $l_1$  est 0.00156 .

Le vecteur solution de la norme  $l_\infty$  est donc

$$\begin{cases} X_1 = 0.99987 \\ X_2 = 0.25353 \\ X_3 = -0.74666 \\ X_4 = 0.24525 \\ X_5 = -0.03751 \end{cases}$$

La valeur de la norme  $l_\infty$  est 0.00130 .

Remarque :

Les tableaux ci-dessus suggèrent que la convergence est quadratique dans les deux cas. Les valeurs du paramètre de contrôle GAM égales à 1 , associées au fait que les vecteurs des multiplicateurs associés aux sous-problèmes linéaires sont strictement positifs, permettent d'ailleurs de croire que les conditions suffisantes de convergence quadratique [chap. II, p. 2.30] sont satisfaites.

(1-2.2) Trouver la meilleure approximation au sens de la norme  $l_1$  et  $l_\infty$  de la fonction

$$\frac{(t + 1)}{(t^2 + 2 * t + 2)}$$

par une somme d'exponentielles de la forme

$$X_1 \exp(X_2 * t) + X_3 \exp(X_4 * t)$$

sur l'ensemble descriptif formé de trente-trois points équipartis dans l'intervalle  $[0, 1]$ , 0 et 1 faisant partie de cet intervalle.

TABLEAUX IV : Approximation de  $(t + 1) / (t^2 + 2 * t + 2)$  par une fonction de la forme  $X_1 \exp(X_2 * t) + X_3 \exp(X_4 * t)$  sur un ensemble de trente-trois points équipartis dans l'intervalle  $[0, 1]$

Norme  $l_\infty$  :

Itération	X				Valeur de la norme	Valeur de GAM
	$X_1$	$X_2$	$X_3$	$X_4$		
	0.5	-0.25	-0.1	-3.	0.1	1.
1	0.54742	-0.31570	-0.04751	-3.31300	0.00315	1.
2	0.55118	-0.31699	-0.05127	-3.50400	0.00013	1.
3	0.55137	-0.31722	-0.05146	-3.48161	0.00009	

Le vecteur solution est donc

$$\begin{cases} X_1 = 0.55137 \\ X_2 = -0.31722 \\ X_3 = -0.05146 \\ X_4 = -3.48161 \end{cases}$$

La valeur de la norme est 0.00009 .

Norme 1 :

Itération	X				Valeur de la norme	Valeur de GAM
	$X_1$	$X_2$	$X_3$	$X_4$		
	0.5	-0.25	-0.1	-3.	1.57561	1.
1	0.54733	-0.31599	-0.04756	-3.32791	0.08587	1.
2	0.55187	-0.31837	-0.05208	-3.48512	0.00189	1.
3	0.55200	-0.31851	-0.05221	-3.46619	0.00163	1.
4	0.55200	-0.31851	-0.05221	-3.46616	0.00163	

Le vecteur solution est donc

$$\begin{cases} X_1 = 0.55200 \\ X_2 = -0.31851 \\ X_3 = -0.05221 \\ X_4 = -3.46616 \end{cases}$$

La valeur de la norme est 0.00163 .

## § 2. NORME EUCLIDIENNE

---

L'algorithme proposé pour résoudre le problème du chapitre IV

$$\min \|f(x)\|_2$$

se compose d'une sous-routine MIN qui permet de résoudre des problèmes

$$\begin{cases} n \leq 35 \\ n > p \\ p \leq 10 \end{cases}$$

où  $n$  est le nombre de composantes de la fonction  $f$  et  $p$  le nombre de variables.

### 2.1. Résultats Numériques.

---

Nous avons utilisé cette sous-routine MIN pour résoudre deux problèmes :

(2-1.1) Approximation d'un vecteur  $y$  formé de valeurs résultant d'une expérience par une fonction de la forme

$$X_1 + X_2 \exp(-X_4 * t) + X_3 \exp(-X_5 * t)$$

sur un ensemble formé de trente-trois valeurs de  $t$ .

Le tableau des données extrait de la référence [3, p. 185] du chapitre IV est le suivant :

i	$t_i$	$y_i$	i	$t_i$	$y_i$	i	$t_i$	$y_i$
1	0	0.844	12	110	0.718	23	220	0.478
2	10	0.908	13	120	0.685	24	230	0.467
3	20	0.932	14	130	0.658	25	240	0.457
4	30	0.936	15	140	0.628	26	250	0.448
5	40	0.925	16	150	0.603	27	260	0.438
6	50	0.908	17	160	0.580	28	270	0.431
7	60	0.881	18	170	0.558	29	280	0.424
8	70	0.850	19	180	0.538	30	290	0.420
9	80	0.818	20	190	0.522	31	300	0.414
10	90	0.784	21	200	0.506	32	310	0.411
11	100	0.751	22	210	0.490	33	320	0.406

Les valeurs choisies pour les paramètres sont

$$\text{EPS} = 1.E-12$$

$$\text{EXP} = 1.5$$

$$\text{DECR} = 0.5$$

$$\text{SIGMA} = 1.E-04 ;$$

GAM est le paramètre assurant la descente de l'algorithme.



TABLEAU V : Approximation au sens des moindres carrés d'un vecteur  $y$  par une fonction de la forme  $X_1 + X_2 \exp(-X_4 * t) + X_3 \exp(-X_5 * t)$

Itération	X					$\ f(x)\ _2^2$	GAM
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$		
	0.5	1.	-1.	0.01	0.01	1.71764	0.25
1	0.28226	1.31460	-0.68539	0.00689	0.01311	0.05821	0.375
2	0.40682	1.23569	-0.75641	0.01067	0.02305	0.03303	0.1875
3	0.36695	1.23535	-0.75885	0.01090	0.02851	0.00120	0.09375
4	0.36019	1.24466	-0.76147	0.01055	0.02813	0.00018	0.04687
5	0.36110	1.26654	-0.78399	0.01068	0.02773	0.00016	0.02343
6	0.36419	1.32908	-0.84902	0.01101	0.02663	0.00015	0.03515
7	0.36664	1.39842	-0.92036	0.01132	0.02574	0.00012	0.01758
8	0.36787	1.44644	-0.96941	0.01151	0.02526	0.000088	0.02637
9	0.36941	1.50593	-1.03012	0.01173	0.02467	0.000083	0.01318
10	0.37028	1.54905	-1.07393	0.01187	0.02433	0.000070	0.01977
11	0.37135	1.60345	-1.12917	0.01204	0.02390	0.000068	0.00988
12	0.37198	1.64345	-1.16966	0.01216	0.02363	0.000062	0.01483
13	0.37276	1.69347	-1.22028	0.01230	0.02331	0.000060	0.00741
14	0.37323	1.73001	-1.25719	0.01239	0.02310	0.000057	0.01112
15	0.37380	1.77432	-1.30194	0.01251	0.02285	0.000057	0.00536
16	0.37415	1.80589	-1.33377	0.01258	0.02270	0.000056	0.00834
17	0.37454	1.84201	-1.37019	0.01266	0.02253	0.000055	0.00417
18	0.37477	1.86643	-1.39479	0.01272	0.02041	0.000055	0.00625
19	0.37502	1.89173	-1.42027	0.01277	0.02230	0.000055	0.00313
20	0.37516	1.90724	-1.43588	0.01280	0.02224	0.000055	0.00313
21	0.37532	1.92606	-1.45484	0.01285	0.02216	0.000055	0.00313
22	0.37538	1.93247	-1.46129	0.01286	0.02213	0.000055	0.00313
23	0.37540	1.93468	-1.46351	0.01286	0.02213	0.000055	



Le vecteur solution est donc

$$\begin{cases} X_1 = 0.37540 \\ X_2 = 1.93468 \\ X_3 = -1.46351 \\ X_4 = 0.01286 \\ X_5 = 0.02213 \end{cases}$$

La valeur de la norme est  $0.5464 \text{ E-04}$  .

Afin de tester une première évaluation des valeurs des paramètres, et en particulier de EXPN et DECR , l'exemple ci-dessus a été testé pour trois valeurs différentes de ce couple de paramètres.

Les résultats obtenus ont été résumés dans le tableau ci-dessous. Rappelons qu'une itération est le passage par la sous-routine qui calcule la matrice gradient en un point donné, et qu'une itération est le passage par la sous-routine qui calcule la valeur de la fonction en un point donné.

Ceci implique que la différence entre le nombre d'itérations intérieures et le nombre d'itérations donne le nombre de fois que nous avons dû accroître le paramètre GAM pour assurer la descente de l'algorithme.

TABLEAU VI : Nombres d'itérations obtenus pour différentes valeurs des paramètres EXPN et DECR dans l'exemple (2-1.1)

EXPN	DECR	Itérations	Itérations Intérieures
1.5	0.5	23	31
3.75	0.25	23	33
10.	0.1	164	244

Au vu de ce tableau, il semble que les valeurs EXPN = 10 et DECR = 0.1 soient peu indiquées. Une analyse des itérations montre d'ailleurs que, pour ces valeurs, la stratégie consistant à diminuer GAM, si la condition de descente est remplie lors de la première résolution du sous-problème linéaire, est peu adéquate car une telle diminution est presque toujours suivie d'un accroissement de GAM à l'itération suivante.

(2-1.2) Approximation au sens des moindres carrés de la fonction

arctang (sin (t))

par un polynôme de la forme

$$X_1 + X_2 t + X_3 t^2 + X_4 t^3 + X_5 t^4 + X_6 t^5$$

sur un ensemble discret formé des points

-1., -0.8, -0.6, -0.4, -0.2, 0., 0.2, 0.4, 0.6, 0.8, 1. .

TABLEAU VII : Approximation au sens des moindres carrés de la fonction  $\arctang(\sin(t))$  par un polynôme du cinquième degré sur un ensemble formé de onze points équi-partis dans l'intervalle  $[-1, 1]$

Itération	X						$\ f(x)\ _2^2$	GAM
	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$		
1	1.	1.	1.	1.	1.	1.	49.78518	0.25
2	-0.00891	0.68616	0.01425	0.13399	0.09177	0.01013	0.08908	0.125
3	-0.00168	0.87422	-0.01943	-0.02896	0.02116	-0.14493	0.00276	0.0625
4	0.00060	0.93121	-0.00480	-0.17819	0.00463	-0.05824	0.00085	0.03125
5	0.43E-04	0.97377	-0.00034	-0.34333	0.00033	0.06772	0.8E-04	0.01562
6	0.8E-06	0.98974	-0.6E-05	-0.40606	0.6E-05	0.11587	0.6E-05	0.00781
7	0.3E-08	0.99137	-0.3E-07	-0.41248	0.3E-07	0.12080	0.6E-05	0.00395
	0.4E-11	0.99141	-0.3E-10	-0.41265	0.3E-10	0.12093	0.6E-05	

Le vecteur solution est donc

$$\begin{cases} X_1 = 0. \\ X_2 = 0.99141 \\ X_3 = 0. \\ X_4 = -0.41265 \\ X_5 = 0. \\ X_6 = 0.12093 \end{cases}$$

La somme des carrés vaut  $0.617E-05$  .

Il est intéressant de noter que l'on obtient un polynôme pair; ceci n'est pas inattendu puisque la fonction  $\arctang(\sin(t))$  est paire et que les points d'interpolation sont opposés par rapport à l'origine.

Notons enfin que, comme pour l'exemple (2-1.1) , nous avons appliqué l'algorithme pour plusieurs valeurs de EXPN et DECR . Les résultats obtenus sont résumés dans le tableau ci-dessous; notons que, dans ce cas-ci, aucun couple de valeurs n'est privilégié.

TABLEAU VIII : Nombres d'itérations obtenus pour différentes valeurs des paramètres EXPN. et DECR dans l'exemple (2-1.2)

EXPN	DECR	Itérations	Itérations intérieures
1.5	0.5	7	7
3.75	0.25	5	5
10.	0.1	5	5

ANNEXE I :

ÉLÉMENTS DE PROGRAMMATION LINÉAIRE

Cette annexe rappellera quelques résultats classiques de programmation linéaire; ces résultats sont bien connus mais leur démonstration réclame souvent certaines connaissances de programmation linéaire dont l'explication serait beaucoup plus longue. Aussi nous contenterons-nous parfois de citer ces résultats en renvoyant le lecteur intéressé par une documentation plus complète au livre de HADLEY [3], par exemple, ou tout autre texte équivalent.

### § 1. DUAL D'UN PROBLEME LINEAIRE

On appelle dual du problème linéaire

$$\left. \begin{array}{l} \min C^T x \quad x(p, 1) \quad C(p, 1) \\ \text{soumis aux contraintes} \\ \begin{array}{l} D \\ (q, p) \end{array} \begin{array}{l} x \\ (p, 1) \end{array} \geq d \\ x_i \geq 0 \quad i = 1, \dots, p \end{array} \right\} \quad (I.1)$$

le problème linéaire suivant :

$$\left. \begin{array}{l} \max [v_1 \quad \dots \quad v_q] d \\ \text{soumis aux contraintes} \\ \begin{array}{l} [v_1 \quad \dots \quad v_q] \end{array} \begin{array}{l} D \\ (q, p) \end{array} \leq \begin{array}{l} C \\ (p, 1) \end{array} \\ v_i \geq 0 \quad i = 1, \dots, q \end{array} \right\} \quad (I.2)$$



A1.2.

On prouve aussi (HADLEY [3, p. 238]) que, si la  $j$ -ième variable du problème primal (I.1) est libre de signe, alors la  $j$ -ième contrainte du dual (I.2) est une égalité.

## § 2. SOLUTION OPTIMALE D'UN PROBLEME LINEAIRE

Il est intéressant de réécrire (I.1) sous la forme

$$\min C^T x$$

sous les contraintes

$$\begin{bmatrix} D & -I_q \\ (p, p+q) \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \\ y_1 \\ \vdots \\ y_q \end{pmatrix} = d$$

$$x_i \geq 0 \quad i = 1, \dots, p$$

$$y_i \geq 0 \quad i = 1, \dots, q$$

où  $I_q$  est la matrice identité de  $\mathbb{R}^q$ .

On montre alors que le programme linéaire admet une solution optimale  $(x^*, y^*)$  si et seulement si

- (1) il existe une sous-matrice  $B(q, q)$  non-singulière de  $[D, -I_q]$  telle que

$$B \begin{pmatrix} x_B^* \\ y_B^* \end{pmatrix} = d$$

où  $[x_B^*, y_B^*]$  désignent les vecteurs correspondant aux colonnes de  $[D, -I_q]$  formant la sous-matrice  $B$  ;

- (2) les autres composantes de  $[x^*, y^*]$  sont nulles.

Voir HADLEY [3, pp. 100-102] .

Notons ici une propriété qui nous intéresse, à savoir

Si  $\text{rg } (D) = p$  et si les variables  $x_i$  sont non-contraintes en signe, alors on peut toujours construire une base  $B(q, q)$  comprenant  $D$  .

Preuve :

La résolution du problème non-constrait en signe par la méthode du simplexe implique la création de nouvelles variables contraintes en signe telles que

$$x_j = x'_j - x''_j \quad j = 1, \dots, q,$$

ce qui permet de construire un tableau simplexe de départ dont la ligne des coûts relatifs est de la forme :

	$x'_1, x''_1, \dots, x'_p, x''_p, y_1, \dots, y_q$
Coûts Relatifs	$\alpha_1, -\alpha_1, \dots, \alpha_p, -\alpha_p, \beta_1, \dots, \beta_q$

La manière dont fonctionne l'algorithme simplexe permet, d'autre part, d'assurer que les coûts relatifs à deux variables  $x'_1$  et  $x''_1$  resteront toujours opposés.

Comme à l'optimum, tous les coûts relatifs deviennent positifs ou nuls; la remarque ci-dessus exige alors que les coûts relatifs des variables  $x'_1$  et  $x''_1$  soient nuls.

Dès lors, ou bien l'une de ces deux variables est en base pour tout  $i$  ,  
ou bien des pivotages sur chaque  $x'_1$  ou  $x''_1$  non en base avec un

élément  $y_j$  en base permettent de créer une base désirée sans changer le coût optimal.

Pour vérifier qu'un tel pivotage est toujours possible, représentons le dernier tableau simplexe en supposant les variables  $(x'_1, \dots, x'_s)$ ,  $s < p$ , en base.

Nous avons un tableau de la forme :

	$x'_1$	$\dots$	$x'_s$	$x'_{s+1}$	$\dots$	$x_p$	$y_1$	$\dots$	$y_q$
$x'_1$	1		0	$\alpha_1$					
$\vdots$									
$x'_s$	0		1	$\alpha_s$					
$y_{i_1}$	0		0	$\alpha_{s+1}$					
$\vdots$									
$y_{i_{q-s}}$	0		0	$\alpha_q$					

Supposer que l'on ne peut pivoter  $x'_{s+1}$  avec une variable  $y_{i_j}$  en base revient à dire que les scalaires  $\alpha_{s+1}$  à  $\alpha_q$  sont nuls; ceci implique que la matrice hachurée est de rang inférieur à  $p$ , ce qui est impossible puisque cette matrice a le même rang que la matrice  $D$  de départ.  $\square$

### § 3. SOLUTION OPTIMALE DUALE

---

Si  $x$  est une solution admissible du problème (I.1) et  $v$  une solution admissible du problème dual (I.2), alors

$$c^T x \geq v^T d$$

Preuve :

Toute solution admissible de (I.2) est telle que

$$(v^T D)_i \leq c_i \quad i = 1, \dots, p \quad (I.3)$$

Comme  $x$  est une solution admissible de (I.1), nous avons

$$x_i \geq 0 \quad i = 1, \dots, p$$

De (I.3), nous tirons donc

$$(v^T D)_i x_i \leq c_i x_i \quad i = 1, \dots, p$$

et, en sommant sur tous les  $i$ , on obtient

$$v^T D x \leq c^T x$$

En partant de (I.1), on prouve semblablement que

$$v^T D x \geq v^T d$$

d'où la thèse.  $\square$

Si  $x^*$  et  $w^*$  sont des solutions admissibles de (I.1) et (I.2) respectivement, on a

$x^*$  et  $w^*$  sont optimales pour (I.1) et (I.2) respectivement si et seulement si

$$c^T x^* = v^{*T} d$$

Nécessité :

Supposons que

$$c^T x^* > v^{*T} d$$

et montrons que nous pouvons construire une autre solution  $v'^*$  admissible pour (I.2) telle que

$$c^T x^* = v'^{*T} d$$

ce qui, compte tenu de l'optimalité de  $v^*$ , est impossible.

Pour cela, supposons que le programme (I.1) se présente sous la forme

$$\left. \begin{array}{l} \min C^T X \\ \text{'soumis aux contraintes} \\ D X = d \quad ; \quad D(q, p) \\ x_i \geq 0 \quad \quad \quad i = 1, \dots, p \\ \text{rg } (D) = q \end{array} \right\} \quad (I.3)$$

Si ce n'est pas le cas, nous ajouterons des variables artificielles mais le



raisonnement est identiquement le même.

Soit  $B$  la base associée à la solution optimale  $x^*$  et  $x_B^*$  le sous-vecteur formé des indices des colonnes de  $D$  entrant en base, les autres composantes de  $x^*$  étant nulles.

Nous savons que

$$x_B^* = B^{-1} d$$

et donc

$$c^T x^* = c_B^T B^{-1} d \quad (I.4)$$

De plus, les coûts relatifs associés à  $x^*$  sont donnés par

$$\bar{c} = c - c^t B^{-1} D$$

le vecteur  $\bar{c}$  étant positif ou nul puisque  $x^*$  est optimal.

Dès lors, en posant

$$v'^{*T} = c B^{-1}$$

nous avons

$$v'^{*T} D = c^t B^{-1} D \leq c$$

ce qui prouve que  $v'^{*}$  est admissible pour (I.2).

De plus,

$$v'^{*T} d = c_B^T B^{-1} d \quad (I.5)$$

(I.4) et (I.5) donnent la contradiction cherchée.  $\square$



Suffisance.

Soit  $x$  une solution admissible pour (I.1) .

Montrons que

$$c^T x \geq c^T x^*$$

Par le théorème ci-dessus, nous avons

$$v^{*T} d \leq c^T x$$

et, par hypothèse,

$$v^{*T} d = c^T x^*$$

ce qui prouve l'assertion.

On montre de la même manière que  $v^*$  est optimale pour (I.2) .  $\square$

#### § 4. THEOREME FAIBLE DES ECARTS COMPLEMENTAIRES

---

Le théorème faible des écarts complémentaires peut s'écrire :

Si  $x^*$  et  $y^*$  sont des solutions optimales des problèmes (I.1) et (I.2) respectivement, nous avons

$$(1) \quad v_i^* - (D x^* - d)_i = 0 \quad i = 1, \dots, q \quad ;$$

$$(2) \quad (c - v^{*T} D)_i x_i^* = 0 \quad i = 1, \dots, p \quad .$$

Preuve :

$x^*$  étant optimal pour (I.1), nous avons

$$D x^* - d \geq 0 \quad . \quad (I.6)$$

$v^*$  étant optimal pour (I.2), nous avons

$$v^* \geq 0$$

Dès lors, (I.6) implique

$$v^{*T} (D x^* - d) \geq 0 \quad . \quad (I.7)$$

On prouve de la même manière que

$$(c - v^{*T} D) x^* \geq 0 \quad . \quad (I.8)$$

Nous savons également qu'à l'optimum,

$$c^T x^* = v^{*T} d$$

Dès lors (I.8) et (I.9) impliquent

$$v^{*T} (D x^* - d) = (c - v^{*T} D) x^* = 0$$

Comme

$$v_i^* \geq 0 \quad i = 1, \dots, q$$

et

$$(D x^* - d)_i \geq 0 \quad i = 1, \dots, q,$$

nous avons

$$v_i^* (D x^* - d)_i = 0 \quad i = 1, \dots, q$$

et semblablement

$$x_i^* (c - v^{*T} D)_i = 0 \quad i = 1, \dots, p \quad .\square$$

ANNEXE II :

Sous-Routines FORTRAN



```

1 SUBROUTINE MINPOL(X,N,P,GNORM,TETA,IPRI,ICRI,SIGMA,EPS,NMAX,NMAXM)
2 C
3 C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
4 C
5 C LA SOUS-ROUTINE MINPOL MINIMISE LA NORME L1 OU DE TCHEBICHEFF
6 C D'UNE FONCTION VECTORIELLE DERIVABLE.
7 C
8 C TABLE DES VARIABLES
9 C *****
10 C
11 C A L'ENTREE:
12 C -----
13 C
14 C X : APPROXIMATION INITIALE.
15 C
16 C N : DIMENSION DE L'ESPACE DE DEPART.
17 C
18 C P : DIMENSION DE L'ESPACE D'ARRIVEE, (N>P).
19 C
20 C TETA : VARIABLE SERVANT A DIMINUER LE PARAMETRE DE CONTROLE GAM.
21 C
22 C IPRI : VARIABLE D'IMPRESSION:
23 C DANS TOUS LES CAS ON IMPRIME: LE VECTEUR INITIAL ET LA VALEUR
24 C DE LA NORME ASSOCIEE, LE VECTEUR
25 C SOLUTION OU LA MEILLEURE APPROXIMATION
26 C OBTENUE, LA VALEUR DE LA NORME ET LA
27 C VALEUR DE LA FONCTION AU POINT X.
28 C
29 C IPRI<0 : IMPRESSIONS MINIMALES.
30 C IPRI>1 : ON IMPRIME TOUTES LES IPRI ITERATIONS LA NOUVELLE
31 C APPROXIMATION ET LA VALEUR DE LA NORME ASSOCIEE.
32 C IPRI=1 : ON IMPRIME A CHAQUE ITERATION LA NOUVELLE APPROXIMATION,
33 C LA VALEUR DE LA NORME ASSOCIEE ET LA VALEUR DU PARAMETRE GAM.
34 C
35 C EPS : VARIABLE SUR LAQUELLE ON EFFECTUE LE TEST DE SORTIE.
36 C
37 C SIGMA : VARIABLE SUR LAQUELLE ON EFFECTUE LE TEST DE DESCENTE,
38 C (0<SIGMA<1)
39 C
40 C NMAX : NOMBRE D'ITERATIONS MAXIMAL AUTORISEES DANS LA SOUS-ROUTINE
41 C DE MINIMISATION LINEAIRE, NMAX<99.
42 C
43 C NMAXM : NOMBRE D'ITERATIONS MAXIMAL AUTORISEES, NMAXM<999.
44 C
45 C ICRI : VARIABLE BOOLEENNE;
46 C ICRI=0 : CAS DE LA NORME DE TCHEBICHEFF
47 C ICRI=1 : CAS DE LA NORME L1
48 C
49 C A LA SORTIE:
50 C -----

```



```

51 C      X :CONTIENT LE VECTEUR SOLUTION OU LA MEILLEURE APPROXIMATION
52 C          OBTENUE.
53 C
54 C      GNORM :CONTIENT LA VALEUR DE LA NORME DE F(X).
55 C
56 C      ICRI :INDIQUE LES TERMINAISONS ANORMALES DE L'ALGORITHME.
57 C      ICRI=-1 :LA MATRICE DES GRADIENTS N'EST PAS DE RANG MAXIMAL.
58 C      ICRI=-2 :LE NOMBRE D'ITERATIONS MAXIMAL ADMIS LORS DE LA
59 C          RESOLUTION DU SOUS-PROBLEME LINEAIRE EST DEPASSE.
60 C
61 C      REMARQUES:
62 C      *****
63 C      1) POUR SON FONCTIONNEMENT LA SOUS-ROUTINE RECLAME L'ADJONCTION
64 C          DE DEUX SOUS-ROUTINES:
65 C          -UNE SOUS-ROUTINE GRAD(A,X) CALCULANT LA MATRICE DES GRADIENTS AU
66 C              POINT X; A :MATRICE DECLAREE (50,25)
67 C              X :VECTEUR.
68 C          -UNE SOUS-ROUTINE FUN(E,X) TELLE QUE E=F(X).
69 C      2) LES DIMENSIONS MAXIMALES DU PROBLEME SONT:NOMBRE DE VARIABLES P < 26
70 C          NOMBRE DE COMPOSANTES DE LA
71 C          FONCTION : N<51.
72 C
73 C      CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
74 C
75 C      IMPLICIT REAL*8(A-H,Q-Z)
76 C      DIMENSION X(1),Y(25),A(50,25),Z(50,50),F(50),H(25)
77 C      INTEGER P
78 C
79 C      INITIALISATIONS
80 C
81 C      KANT=0
82 C      KRI=0
83 C      IPRO=-1
84 C      CALL FUN(F,X)
85 C      CALL NORM(GNORM1,ICRI,F,N)
86 C      INIT=1
87 C      ICRC=ICRI
88 C
89 C      IMPRESSIONS INITIALES
90 C
91 C      PRINT 9000
92 C      CALL PRI(X,P)
93 C      PRINT 9109
94 C      CALL PRI(F,N)
95 C      PRINT 9001,GNORM1
96 C
97 C      KANT=KANT+1
98 C      KRI=KRI+1
99 C
100 C      IMPRESSIONS EVENTUELLES

```

```

101 C      IF(IPRI.LT.0) GOTO 41
102      IF(IPRI.NE.1) GOTO 604
103      PRINT 9005,KANT
104      PRINT 9010
105      GOTO 605
106      IF(KRI.LT.IPRI) GOTO 41
107 604    PRINT 9005,KANT
108      PRINT 9010
109      PRINT 9008
110 605    CALL PRI(X,P)
111      PRINT 9001,GNORM1
112      KRI=0
113
114 C
115 C      CALCUL DE LA MATRICE DES GRADIENTS AU POINT X
116 C
117 41     CALL GRAD(A,X)
118      GAM=1.
119 C
120 C      RESOLUTION DU SOUS-PROBLEME LINEAIRE:MINIMISER LA NORME. DE
121 C      F(X)+A(X)H (H VARIABLE).
122 C
123      CALL POL(A,N,P,F,ICRO,IPRO,NMAX,H,SUM,INIT,Z)
124      IF(ICRO.EQ.-2.OR.ICRO.EQ.-1) RETURN
125      IF(INIT.NE.-1) GOTO 5
126      INIT=1
127      ICRO=ICRI
128      CALL POL(A,N,P,F,ICRO,IPRO,NVAX,H,SUM,INIT,Z)
129      IF(ICRO.EQ.-2.OR.ICRO.EQ.-1) RETURN
130 C
131 C      CONSTRUCTION DE LA NOUVELLE APPROXIMATION Y=X+GAM*H.
132 C
133 5       DO 1 I=1,P
134 1       Y(I)=X(I)+GAM*H(I)
135      IF(IPRI.NE.1) GOTO 606
136      PRINT 9011,GAM
137 606    CALL FUN(F,Y)
138      CALL NORM(GNORM2,ICRI,F,N)
139      T=GNORM1-SUM
140 C
141 C      TEST D'OPTIMALITE DE X.
142 C
143      IF(T.LE.EPS) GOTO 10
144      IF(((GNORM1-GNORM2)/(T*GAM)).LE.SIGMA) GOTO 3
145      DO 2 I=1,P
146 2       X(I)=Y(I)
147      GNORM1=GNORM2
148      INIT=0
149 C
150 C      TEST DETERMINANT SI LE NOMBRE MAXIMAL D'ITERATIONS EST ATTEINT.

```



```

151 C
152 IF(KANT.GT.NMAXM) GOTO 11
153 GOTO 4
154 C
155 C DIMINUTION DU PARAMETRE GAM
156 C
157 3 GAM=GAM*TETA
158 GOTO 5
159 C
160 C IMPRESSIONS FINALES.
161 C
162 11 PRINT 9003,NMAXM
163 PRINT 9004
164 CALL PRI(X,P)
165 PRINT 9009
166 CALL PRI(F,N)
167 PRINT 9001,GNORM1
168 RETURN
169 10 PRINT 9006
170 CALL PRI(X,P)
171 PRINT 9009
172 CALL PRI(F,N)
173 GNORM=GNORM1
174 PRINT 9001,GNORM
175 C
176 9000 FORMAT(1H1,'VECTEUR INITIAL:',/)
177 9001 FORMAT(1X,'VALEUR DE LA NORME:',D16.8)
178 9003 FORMAT(1H1,'LE NOMBRE D ITERATIONS MAXIMAL'13'EST DEPASSE')
179 9004 FORMAT(1X,/,1X,'LA MLILLEURE APPROXIMATION OBTENUE EST:')
180 9005 FORMAT(1X,/,1X,'ITERATION:',13,/)
181 9006 FORMAT(1H1,'VECTEUR SOLUTION:')
182 9008 FORMAT(1X,'NOUVELLE APPROXIMATION:')
183 9009 FORMAT(1X,'VALEUR DE LA FONCTION:')
184 9010 FORMAT(1X,'*****',/)
185 9011 FORMAT(/,1X,'VALEUR DU PARAMETRE DE CONTROLE GAM:'D16.8,/)
186 13 RETURN
187 END

```

```

1      SUBROUTINE NORM(GNORM,ICRI,X,N)
2 C
3 C      LA SOUS-ROUTINE NORM DETERMINE LA NORME L1 OU DE TCHEBICHEFF D'UN
4 C      VECTEUR X DONNE.
5 C
6      IMPLICIT REAL *8(A-H,Q-Z)
7      DIMENSION X(1)
8      IF(ICRI)1,2,1
9      1      GNORM=0.
10     DO 3 I=1,N
11     3      GNORM=GNORM+DABS(X(I))
12     RETURN
13     2      GNORM=0.
14     DO 4 I=1,N
15     4      IF(GNORM.GE.DABS(X(I))) GOTO 4
16     GNORM=DABS(X(I))
17     4      CONTINUE
18     RETURN
19     END

```



```

1 SUBROUTINE POL(A,N,P,F,ICRI,IPRI,NMAX,H,SUM,INIT,Z)
2 C
3 C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
4 C
5 C LA SOUS-ROUTINE POL DETERMINE LE VECTEUR H MINIMISANT LA NORME L1 OU DE
6 C TCHEBICHEFF DU VECTEUR F+AH
7 C
8 C TABLE DES VARIABLES
9 C *****
10 C A L'ENTREE:
11 C -----
12 C
13 C A:MATRICE DES VARIABLES APPROXIMANTES.
14 C
15 C F:VECTEUR A APPROXIMER.
16 C
17 C ICRI:VARIABLE BOOLEENNE
18 C ICRI=1:CAS DE LA NORME L1
19 C ICRI=0:CAS DE LA NORME DE TCHEBICHEFF
20 C
21 C IPRI:VARIABLE D'IMPRESSION
22 C IPRI<0,AUCUNE IMPRESSION
23 C IPRI=0,ON IMPRIME LA MATRICE A DU PROBLEME,LE VECTEUR F A APPROXIMER
24 C LE VECTEUR SOLUTION ET LA VALEUR DE LA NORME CORRESPONDANTE
25 C IPRI>0,ON IMPRIME LA MATRICE A DU PROBLEME,LE VECTEUR F A APPROXIMER
26 C LE VECTEUR SOLUTION , LA VALEUR DE LA NORME CORRESPONDANTE
27 C ET A CHAQUE ITERATION LA REFERENCE UTILISEE,LA NOUVELLE
28 C APPROXIMATION ET LA VALEUR DE LA NORME CORRESPONDANTE.
29 C
30 C NMAX:NOMBRE MAXIMAL D'ITERATIONS,NMAX<99.
31 C
32 C INIT:VARIABLE INDIQUANT SI LA REFERENCE INITIALE EST DONNEE OU NON
33 C INIT=1:LA REFERENCE INITIALE EST DONNEE.
34 C INIT=0:LA REFERENCE INITIALE N'EST PAS DONNEE.
35 C
36 C Z:MATRICE QUI,DANS LE CAS OU LA REFERENCE INITIALE EST DONNEE,CONTIENT
37 C LES (P+1) LIGNES DE CETTE REFERENCE.
38 C
39 C N ET P:DIMENSIONS DE LA MATRICE A(N,P)
40 C N>P
41 C N<51
42 C P<26
43 C
44 C A LA SORTIE:
45 C -----
46 C
47 C H:VECTEUR SOLUTION
48 C
49 C SUM:VALEUR DE LA NORME CORRESPONDANTE
50 C

```

```

51 C      ICRI ET INIT INDIQUENT LES TYPES DE TERMINAISONS ANORMALES DE
52 C      L'ALGORITHME
53 C      ICRI=-1:LA MATRICE A N'EST PAS DE RANG MAXIMAL
54 C      ICRI=-2:LE NOMBRE MAXIMAL D'ITERATIONS EST DEPASSE
55 C      INIT=-1:DANS LE CAS OU LA REFERENCE INITIALE EST DONNEE, CELLE-CI
56 C      NE CONDUIT PAS A UNE MATRICE Z*A DE RANG MAXIMAL.
57 C
58 C      Z :CONTIENT LA REFERENCE OPTIMALE
59 C
60 C      CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
61 C
62 C      IMPLICIT INTEGER (P)
63 C      IMPLICIT REAL*8(A-H,Q-Z)
64 C      DIMENSION B(50,50),A(50,25),Z(50,50),ZA(50,25),ZO(50,27),O(50,27)
65 C      DIMENSION U(50,27),F(1),V(27),UL(25),T(50),S(50)
66 C      DIMENSION SA(50),H(1),W(27)
67 C
68 C      INITIALISATIONS
69 C
70 C      PA=P+1
71 C      PI=P+2
72 C      EF=0.00000001
73 C      KANT=0
74 C      X=1.
75 C      GPI=4.*DATAN(X)
76 C
77 C      APPEL A LA SOUS-ROUTINE AUX QUI FOURNIT LES REFERENCES ET VECTEUR
78 C      COMPLEMENTAIRE INITIAUX DE L'ALGORITHME.
79 C
80 C      CALL AUX(A,N,P,Z,ICRI,V,INIT)
81 C
82 C      IF(INIT.LT.0) RETURN
83 C
84 C      IMPRESSIONS EVENTUELLES
85 C
86 C      IF(ICRI) 67,17,7
87 C 67 PRINT 9010
88 C      PRINT 9012
89 C      PRINT 9011
90 C      CALL PRO(A,N,P)
91 C      GOTO 101
92 C
93 C      CONSTRUCTION DE LA MATRICE B DEFINISSANT LA NORME POLYHEDRALE
94 C      EN FONCTION DE LA NORME CHOISIE.
95 C
96 C 17 DO 27 I=1,N
97 C      DO 27 J=1,N
98 C      IF(I-J) 37,47,37
99 C 37 B(I,J)=0.
100 C      GOTO 27

```



```

101 47 B(I,J)=1.
102 27 CONTINUE
103 MI=N
104 GOTO 281
105 7 DO 57 I=1,N
106 57 R(1,I)=1.
107 MI=1
108 281 IF(IPRI.LT.0) GOTO 280
109 PRINT 9010
110 PRINT 9000
111 CALL PRO(A,N,P)
112 PRINT 9001
113 CALL PRI(F,N)
114 C
115 C ESTIMATION DE LA NORME
116 C
117 280 SUM=0.
118 V(PI)=0.
119 KANT=KANT+1
120 DO 20 I=1,PA
121 R=0.
122 DO 30 J=1,N
123 30 R=R+Z(I,J)*F(J)
124 20 SUM=SUM+R*V(I)
125 IF(SUM.GE.-EP) GOTO 50
126 DO 40 I=1,PA
127 40 DO 40 J=1,N
128 40 Z(I,J)=-Z(I,J)
129 SUM=-SUM
130 50 CALL PROD(Z,A,ZA,PA,N,P)
131 C
132 C IMPRESSIONS EVENTUELLES
133 C
134 IF(IPRI.LE.0) GOTO 64
135 PRINT 9002,KANT
136 PRINT 9003
137 PRINT 9000,SUM
138 PRINT 9004
139 CALL PRO(Z,PA,N)
140 64 CALL PRMAVC(Z,F,T,PA,N)
141 C
142 C FACTORISATION DE LA MATRICE ZA SOUS LA FORME QU ;U TRIANGULAIRE
143 C SUPERIEURE,Q ORTHOGONALE.
144 C
145 CALL HAUS(ZA,Q,U,PA,P)
146 DO 60 I=1,PA
147 60 DO 63 J=1,PA
148 63 ZO(I,J)=Q(J,I)
149 C
150 C CALCUL DU VECTEUR H CANDIDAT A L'OPTIMALITE

```

```

151 C
152 60 SA(I)=SUM-T(I)
153 CALL PRMAVC(ZO,SA,T,PA,PA)
154 H(P)=T(P)/U(P,P)
155 PU=P-1
156 DO 41 I=1,PU
157 R=0.
158 IR=P-I+1
159 DO 51 J=IR,P
160 51 R=R+H(J)*U(IR-1,J)
161 41 H(P-I)=(T(P-I)-R)/U(P-I,P-I)
162 IF(IPRI.LT.1) GOTO 74
163 PRINT 9006
164 CALL PRI(H,P)
165 C
166 C TEST D'OPTIMALITE DE H OU DETERMINATION DE LA LIGNE ENTRANT
167 C DANS LA NOUVELLE REFERENCE
168 C
169 74 CALL PRMAVC(A,H,SA,N,P)
170 DO 70 I=1,N
171 70 T(I)=SA(I)+F(I)
172 GMAX=0.
173 DO 80 I=1,MI
174 R=0.
175 DO 90 J=1,N
176 90 K=R+B(I,J)*DABS(T(J))
177 IF(GMAX.GT.(R-SUM)) GOTO 80
178 GMAX=R-SUM
179 KI=I
180 80 CONTINUE
181 C
182 C TEST D'OPTIMALITE
183 IF(GMAX.LE.EP) GOTO 100
184 INIT=1
185 C
186 C DETERMINATION DE LA LIGNE ENTRANT DANS LA REFERENCE SUIVANTE
187 C
188 DO 121 J=1,P
189 R=0.
190 DO 122 I=1,N
191 RA=1.
192 RA=SIGN(RA,T(I))
193 122 P=R+RA*B(KI,I)*A(I,J)
194 121 SA(J)=R
195 DO 123 I=1,N
196 KA=1.
197 RA=SIGN(RA,T(I))
198 123 UL(I)=RA*B(KI,I)
199 DO 110 I=1,PA
200 G(P,I)=0.

```



```

201 110 Q(I,PI)=0.
202      Q(PI,PI)=1.
203 C
204 C      CALCUL PAR TRIANGULARISATION DU VECTEUR T TEL QUE SA=TU
205 C
206      T(1)=SA(1)/U(1,1)
207      DO 71 I=2,P
208          R=0.
209          IR=I-1
210          DO 81 J=1,IR
211 81      R=R+T(J)*U(J,I)
212 71      T(I)=(SA(I)-R)/U(I,I)
213 C
214 C      I C
215 C      FACTORISATION DE LA MATRICE Q SOUS LA FORME P Q C TRIANGULAIRE
216 C      T O
217 C      SUPERIEURE, P PRODUIT DE MATRICES DE ROTATIONS
218 C
219      DO 130 I=1,PI
220 130      S(I)=1.
221          S(P+1)=0.
222          DO 120 I=1,P
223              IF(DABS(T(I)).GT.EP) GOTO 133
224              S(I)=0.
225              GOTO 120
226 133      TETA=DATAN(-T(I))
227 C
228 C      TEST AFIN D'ASSURER QUE W(P+2)>0
229          IF(DCOS(TETA).LE.-EP) TETA=TETA+GPI
230          TCOS=DCOS(TETA)
231          IF(I.EQ.P) GOTO 170
232          IR=I+1
233          DO 140 J=IR,P
234 140      T(J)=T(J)*TCOS
235 170      IF(I.EQ.1) GOTO 180
236          IR=I-1
237 C
238 C      Q0 0
239 C      CALCUL DU VECTEUR S TEL QUE S= P
240 C      01 1
241      DO 190 J=1,IR
242 190      S(J)=S(J)*TCOS
243 180      S(I)=DSIN(TETA)*S(I)
244          S(PI)=TCOS*S(PI)
245 120      CONTINUE
246          CALL PRMAVC(Q,S,W,PI,PI)
247
248 C      DETERMINATION DE LA LIGNE SORTANT DE LA REFERENCE
249 C
250      GMIN=0.

```

```

251      DO 220 I=1,PA
252      IF(V(I).LE.EP) GOTO 220
253      GMIN=W(I)/V(I)
254      KI=I
255      GOTO 241
256 220   CONTINUE
257 241   DO 200 I=1,PA
258      IF(V(I).GT.EP) GOTO 210
259      IF(W(I).GE.-EP) GOTO 200
260      KI=I
261      GOTO 221
262 210   IF((W(I)/V(I)).GT.GMIN) GOTO 200
263      GMIN=W(I)/V(I)
264      KI=I
265 200   CONTINUE
266 C
267 C      CONSTRUCTION DU NOUVEAU VECTEUR DES MULTIPLICATEURS
268 C
269 221   R=0.
270      RA=W(KI)
271      RO=W(KI)
272      DO 250 I=1,PI
273      V(I)=RA*W(I)-RO*V(I)
274 250   R=R+V(I)
275      V(KI)=V(PI)
276      DO 260 I=1,PA
277 260   V(I)=V(I)/R
278 C
279 C      CONSTRUCTION DE LA NOUVELLE REFERENCE
280 C
281      DO 270 J=1,N
282 270   Z(KI,J)=UL(J)
283 C
284 C      TEST DETERMINANT SI LE NOMBRE D'ITERATIONS EST INFERIEUR AU NOMBRE
285 C      MAXIMAL:NMAX
286 C
287      IF(KANT.LT.NMAX) GOTO 280
288      PRINT 9011,NMAX
289      ICPI=-2
290      GOTO 101
291 C
292 C      IMPRESSIONS FINALES
293 C
294 100   IF(IPRI.LT.0) GOTO 101
295      PRINT 9010
296      PRINT 9008,SUM
297      PRINT 9007
298      CALL PRICH,P)
299      IF(IPRI.LE.0) GOTO 101
300      PRINT 9009

```



```
301      CALL PRO(Z,PA,N)
302 c
303 9000 FORMAT(1X,'MATRICE DU PROBLEME:')
304 9001 FORMAT(1X,'VECTEUR A APPROXIMER:')
305 9002 FORMAT(1X,'ITERATION:',1X,I2)
306 9003 FORMAT(1X,/,1X,'*****')
307 9004 FORMAT(1X,'REFERENCE UTILISEE:')
308 9006 FORMAT(1X,'NOUVAU CANDIDAT A L OPTIMALITE:')
309 9007 FORMAT(1X,'VECTEUR SOLUTION:')
310 9008 FORMAT(1X,'VALEUR DE LA NORME:',D16.8,/)
311 9009 FORMAT(1X,/,1X,'REFERENCE OPTIMALE:')
312 9010 FORMAT(1H1)
313 9011 FORMAT(1X,/,1X,'LE NOMBRE D ITERATIONS MAXIMAL'I3'EST DEPASSE')
314 9012 FORMAT(1X,'LA MATRICE DU PROBLEME N EST PAS DE RANG MAXIMAL')
315 101  RETURN
316      END
```



```

1      SUBROUTINE PRO(A,N,M)
2 C
3 C      LA SOUS-ROUTINE PRO EST UNE SOUS-ROUTINE D'IMPRESSIONS MATRICIELLES.
4 C
5      IMPLICIT REAL*8(A-H,Q-Z)
6      DIMENSION A(50,1)
7      PRINT 9000
8      DO 1 I=1,N
9      PRINT 9002
10     1 PRINT 9001,(A(I,J),J=1,M)
11     PRINT 9000
12     9000 FORMAT(1X,/)
13     9001 FORMAT(1X,7D16.8)
14     9002 FORMAT(1X,/)
15     RETURN
16     END

```

```

1      SUBROUTINE PRI(A,N)
2 C
3 C      LA SOUS-ROUTINE PRI EST UNE SOUS-ROUTINE D'IMPRESSIONS VECTORIELLES.
4 C
5      IMPLICIT REAL*8(A-H,Q-Z)
6      DIMENSION A(1)
7      PRINT 9000
8      PRINT 9001,(A(I),I=1,N)
9      PRINT 9000
10     9000 FORMAT(1X,/)
11     9001 FORMAT(1X,7D16.8)
12     RETURN
13     END

```



```

1      SUBROUTINE PROD(A,B,C,M,N,P)
2 C
3 C      LA SOUS-ROUTINE PROD EFFECTUE LE PRODUIT DE DEUX MATRICES A(M,N) ET
4 C      B(N,P).
5 C
6      IMPLICIT REAL*8(A-H,Q-Z)
7      DIMENSION A(50,1),B(50,1),C(50,1)
8      INTEGER P
9      DO 1 I=1,M
10     DO 2 J=1,P
11     RS=0.
12     DO 3 K=1,N
13     3   RS=RS+A(I,K)*B(K,J)
14     2   C(I,J)=RS
15     1   CONTINUE
16     RETURN
17     END

```

```

1      SUBROUTINE PRMAVC(A,F,H,M,N)
2 C
3 C      LA SOUS-ROUTINE PRMAVC EFFECTUE LE PRODUIT D'UNE MATRICE A(M,N)
4 C      ET D'UN VECTEUR F(N).
5 C
6      IMPLICIT REAL*8(A-H,Q-Z)
7      DIMENSION A(50,1),F(1),H(1)
8      DO 1 I=1,M
9      R=0.
10     DO 2 J=1,N
11     2   R=R+A(I,J)*F(J)
12     1   H(I)=R
13     RETURN
14     END

```



```

1      SUBROUTINE HAUS(A,W,S,N,P)
2 C
3 C      LA SOUS-ROUTINE HAUS FACTORISE UNE MATRICE A(N,P) (RG A=P) SOUS LA
4 C      FORME: A=WS, W(N,N) ORTHOGONALE, S(N,P), TRIANGULAIRE SUPERIEURE.
5 C
6      INTEGER P
7      IMPLICIT REAL*8(A-H,Q-Z)
8      DIMENSION A(50,27), W(50,50), S(50,27)
9      DIMENSION V(50), E(50)
10     DO 2 I=1,N
11     DO 2 J=1,P
12 2     S(I,J)=A(I,J)
13     DO 19 I=1,N
14     DO 19 J=1,N
15     IF(I-J) 20,21,20
16 20     W(I,J)=0.
17     GOTO 19
18 21     W(I,J)=1.
19 19     CONTINUE
20     DO 3 KI=1,P
21     T=0.
22     DO 4 I=KI,N
23 4     T=T+S(I,KI)**2.
24     D=SQRT(T)
25     IF((D-DABS(S(KI,KI))) .LE. D.000001) GOTO 3
26     D=-SIGN(D,S(KI,KI))
27     Y=2.*D*(D-S(KI,KI))
28     DO 5 J=KI,N
29 5     V(J)=S(J,KI)
30     V(KI)=V(KI)-D
31     DO 6 J=KI,N
32     LO 17 K=KI,N
33 17     E(K)=S(K,J)
34     DO 6 I=KI,N
35     R=0.
36     DO 7 K=KI,N
37 7     R=R-2.*V(I)*V(K)*E(K)/Y
38     R=R+S(I,J)
39 6     S(I,J)=R
40     DO 22 I=1,N
41     DO 18 K=1,N
42 18     E(K)=W(I,K)
43     DO 22 J=KI,N
44     R=0.
45     DO 23 K=KI,N
46 23     R=R-2.*E(K)*V(K)*V(J)/Y
47 22     W(I,J)=R+W(I,J)
48 3     CONTINUE
49     RETURN
50     END

```



```

1 SUBROUTINE AUX(C,N,P,Z,ICRI,V,INIT)
2 C
3 C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
4 C
5 C LA SOUS-ROUTINE AUX DETERMINE LES QUANTITES INITIALES NECESSAIRES A LA
6 C SOUS-ROUTINE POL :A SAVOIR
7 C UNE REFERENCE Z FORMEE DE (P+1) LIGNES (EVENTUELLEMENT FOURNIE A
8 C L'ENTREE DE LA SOUS-ROUTINE) TELLE QUE LA MATRICE Z*C SOIT DE RANG
9 C MAXIMAL : P.
10 C UN VECTEUR V DES MULTIPLCATEURS TEL QUE: V*ZC=0
11 C V*V=1
12 C V POSITIF OU NUL.
13 C
14 C TABLE DES VARIABLES
15 C *****
16 C A L'ENTREE:
17 C -----
18 C
19 C C :MATRICE DES VARIABLES APPROXIMANTES DE LA SOUS-ROUTINE POL.
20 C
21 C Z :CONTIENT LES P+1 LIGNES DE LA REFFRENCE INITIALE SI CELLE-CI EST
22 C DONNEE.
23 C
24 C ICRI :VARIABLE BOOLEENNE
25 C ICRI=0: CAS DE LA NORME DE TCHEBICHEFF
26 C ICRI=1: CAS DE LA NORME L1.
27 C
28 C INIT: VARIABLE INDICANT SI LA REFERENCE EST DONNEE OU NON.
29 C INIT=0: LA REFERENCE EST DONNEE
30 C INIT=1: LA REFERENCE EST A DETERMINER.
31 C
32 C N ET P SONT LES DIMENSIONS DE LA MATRICE C(N,P) (N>P).
33 C
34 C A LA SORTIE:
35 C -----
36 C
37 C Z :CONTIENT LA REFERENCE CHERCHEE.
38 C
39 C V :VECTEUR DES MULTIPLICATEURS.
40 C
41 C INIT ET ICRI INDICENT LES TERMINAISONS ANORMALES DE LA SOUS-ROUTINE
42 C ICRI=-1: LA MATRICE C N'EST PAS DE RANG MAXIMAL:
43 C INIT=-1: DANS LE CAS OU LA REFERENCE EST DONNEE INITIALEMENT INDIQUE
44 C QUE CELLE-CI NE CONDUIT PAS A UNE MATRICE Z*C DE RANG MAXIMAL.
45 C
46 C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
47 C
48 IMPLICIT REAL*8(A-H,O-Z)
49 INTEGER P,PI,PU,PA
50 DIMENSION C(50,25),Z(50,50),V(25),INT2(25),INT(25),D(50,25),

```

```

51      CZA(50,25)
52      PI=P+1
53      IR=0
54      DO 1 I=1,P
55      INT2(I)=0
56 1     INT(I)=I
57      KANT=0
58      KINT=1
59 C
60 C      INITIALISATIONS DE LA MATRICE Z EN FONCTION DE LA NORME UTILISEE.
61 C
62      IF(INIT.EQ.0) GOTO 24
63      IF(ICRI.EQ.0) GOTO 20
64      DO 12 I=1,N
65      DO 12 J=1,N
66      IF(I-J) 13,14,13
67 13     Z(I,J)=1.
68      GOTO 12
69 14     Z(I,J)=-1.
70 12     CONTINUE
71      Z(1,1)=1.
72      CALL PROD(Z,C,D,N,N,P)
73      GOTO 1J
74 20     DO 21 I=1,N
75      DO 21 J=1,N
76      IF(I-J)22,23,22
77 22     Z(I,J)=0.
78      GOTO 21
79 23     Z(I,J)=1.
80 21     D(I,J)=C(I,J)
81 C
82 C      DETERMINATION PAR PIVOTAGE DE P LIGNES DE LA MATRICE ZC LINEAIREMENT
83 C      INDEPENDANTES.
84 C
85 10     KRI=P-KANT
86      DO 2 I=KINT,N
87      DO 2 J=1,KRI
88      K=INT(J)
89      IF(DABS(D(I,K)).LE.0.000001) GOTO2
90      KI=I
91      KJ=K
92      KO=J
93      IF(KI.NE.KINT) IR=KINT
94      GOTO 3
95 2      CONTINUE
96      GOTO 11
97 3      INT2(KANT+1)=KI
98      IF(KANT.EQ.(P-1)) GOTO 50
99      KFO=KI+1
100     R=D(KI,KJ)

```



```

101      DO 5 I=KRO,N
102      DO 5 J=1,KRI
103      K=INT(J)
104      IF(K.EQ.KJ) GOTO 5
105      D(I,K)=D(I,K)-D(I,KJ)*D(KI,K)/R
106 5      CONTINUE
107      INT(KO)=INT(KRI)
108      KANT=KANT+1
109      KINT=KI+1
110      GOTO 10
111 11      ICPI=-1
112      GOTO 52
113 50      IF(IR.EQ.0) IR=P+1
114 C
115 C      DETERMINATION PAR GAUSS DU VECTEUR V TEL QUE V*ZC=0 OU Z CONTIENT
116 C      LES P+1 LIGNES DETERMINEES CI-AVANT.
117 C
118      DO 16 I=1,P
119      K=INT2(I)
120      DO 16 J=1,N
121 16      ZA(I,J)=Z(K,J)
122      DO 32 I=1,N
123 32      ZA(PI,I)=Z(IR,I)
124 51      DO 33 I=1,PI
125      DO 33 J=1,N
126 33      Z(I,J)=ZA(I,J)
127 24      CALL PROD(Z,C,ZA,PI,N,P)
128      PU=P-1
129      PA=P+1
130      DO 125 I=1,P
131      DO 115 J=1,P
132 115      D(J,I)=ZA(I,J)
133 125      D(I,PA)=-ZA(PA,I)
134      DO 175 I=1,PU
135      IO=I+1
136      IF(DABS(D(I,IO)).GT.0.0000001) GOTO 25
137      DO 35 K=IO,P
138      IF(DABS(D(K,IO)).LE.0.000001) GOTO 35
139      KI=K
140      COTO 45
141 35      CONTINUE
142      INIT=-1
143      RETURN
144 45      DO 55 K=I,P
145      DI=D(I,K)
146      D(I,K)=D(KI,K)
147 55      D(KI,K)=DI
148      DI=D(I,PA)
149      D(I,PA)=D(KI,PA)
150      D(KI,PA)=DI

```

```

151 25  QI=D(I,I)
152      DO 65 K=I,P
153 65  D(I,K)=D(I,K)/QI
154      D(I,PA)=D(I,PA)/QI
155      DO 75 L=I0,P
156      DO 85 K=I0,P
157 85  D(L,K)=D(L,K)-D(L,I)*D(I,K)
158 75  D(L,PA)=D(L,PA)-D(I,PA)*D(L,I)
159 175  CONTINUE
160      IF(DABS(D(P,P)).GE.0.000001) GOTO 176
161      INIT=-1
162      RETURN
163 176  V(P)=D(P,PA)/D(P,P)
164      DO 95 I=1,PJ
165      R=C0
166      IR=P-I+1
167      DO 105 J=IR,P
168 105  R=R+V(J)*D(IR-1,J)
169 95  V(P-I)=(D(P-I,PA)-R)/D(P-I,P-I)
170 C
171 C  MODIFICATIONS EVENTUELLES DE LA REFERENCE POUR QUE LE VECTEUR V SOIT
172 C  POSITIF.
173 C
174      R=00
175      DO 135 I=1,P
176      R=R+DABS(V(I))
177      IF(V(I).GE.00) GOTO 135
178      DO 145 K=1,N
179 145  Z(I,K)=-Z(I,K)
180      V(I)=-V(I)
181 135  CONTINUE
182      V(PA)=10
183      R=R+10
184 C
185 C  NORMALISATION DE V.
186 C
187      DO 155 I=1,PA
188 155  V(I)=V(I)/R
189 52  RETURN
190      END

```



```

1      SUBROUTINE MIN(X,N,P,SIGMA,GAM,EXPN,GMAX,DECR,EPS,NMAX,IPRI,SUM1)
2 C
3 C      CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
4 C
5 C      LA SOUS-ROUTINE MIN DETERMINE LES POINTS STATIONNAIRES DE LA,
6 C      NORME EUCLIDIENNE D'UNE FONCTION DERIVABLE
7 C
8 C      TABLE DES VARIABLES
9 C      *****
10 C
11 C      A L'ENTREE :
12 C      -----
13 C      X: APPROXIMATION INITIALE
14 C
15 C      N: DIMENSION DE L'ESPACE D'ARRIVEE
16 C
17 C      P: DIMENSION DE L'ESPACE DE DEPART.
18 C      (N> OU =P)
19 C
20 C      GAM: PARAMETRE DE LA MATRICE AUGMENTEE
21 C
22 C      SIGMA: NOMBRE SUR LEQUEL ON EFFECTUE LE TEST
23 C      DE DESCENTE
24 C
25 C      DECR: VARIABLE SERVANT A DIMINUER GAM
26 C
27 C      EXPN: VARIABLE SERVANT A ACCROITRE GAM
28 C
29 C      EPS: VARIABLE SUR LAQUELLE ON EFFECTUE LE TEST
30 C      D'ARRET
31 C
32 C      GMAX: VARIABLE CONTROLANT LA DIMENSION DE GAM (GMAX>100)
33 C
34 C      NMAX: NOMBRE MAXIMAL D'ITERATIONS AUTORISEES (NMAX<1000)
35 C
36 C      IPRI: VARIABLE D'IMPRESSION
37 C      DANS TCUS LES CAS ON IMPRIME:
38 C      LE VECTEUR INITIAL
39 C      LE VECTEUR SOLUTION
40 C      AVEC LA VALEUR DE LA SOMME DES CARRES
41 C      LE NOMBRE D' ITERATIONS ET LE NOMBRE
42 C      D'ITERATIONS INTERIEURES
43 C      IPRI<0: IMPRESSIONS MINIMALES C'EST-A-DIRE
44 C      IMPRESSIONS CI-DESSUS
45 C      IPRI=0: ON IMPRIME LA MATRICE DES GRADIENTS ET
46 C      LA VALEUR DE LA FONCTION AU POINT OPTIMAL
47 C      IPRI>1: ON IMPRIME TOUTES LES IPRI ITERATIONS
48 C      LA NOUVELLE APPROXIMATION ET LA SOMME DES
49 C      CARRES CORRESPONDANTE AINSI QUE LA MATRICE DES
50 C      GRADIENTS ET LA VALEUR DE LA FONCTION AU

```

```

51 C
52 C
53 C
54 C
55 C
56 C
57 C
58 C
59 C
60 C
61 C
62 C
63 C
64 C
65 C
66 C
67 C
68 C
69 C
70 C
71 C
72 C
73 C
74 C
75 C
76 C
77 C
78 C
79 C
80 C
81 C
82 C
83 C
84 C
85 C
86 C
87 C
88 C
89 C
90 C
91 C
92 C
93 C
94 C
95 C
96 C
97 C
98 C
99 C
100 C

```

PCINT OPTIMAL  
 IPRI=1: MEMES IMPRESSIONS QUE IPRI>1 MAIS A  
 CHAQUE ITERATION INTERIEURE.  
 ON Y AJOUTE EGALEMENT LA VALEUR DU PARAMETRE  
 DE CONTROLE.

A LA SORTIE :  
 -----

X: VECTEUR SOLUTION OU MEILLEURE APPROXIMATION OBTENUE  
 SUM1: SOMME DES CARRES CORRESPONDANTE

REMARQUES :  
 \*\*\*\*\*

1) LES DIMENSIONS MAXIMALES DU PROBLEME SONT : P<11  
 N<36

2) POUR SON FONCTIONNEMENT LA SOUS-ROUTINE MIN RECLAME L'AJONCTION  
 DE DEUX SOUS-ROUTINES :  
 - UNE SOUS-ROUTINE GRAD(S,X) CALCULANT LA MATRICE DES GRADIENTS  
 AU POINT X: S MATRICE DECLAREE (45,10)  
 X VECTEUR  
 - UNE SOUS-ROUTINE FUN(E,X) TELLE QUE E=F(X).

3) NOUS APPELONS ITERATION TOUT PASSAGE PAR LA SOUS-ROUTINE GRAD ET  
 ITERATION INTERIEURE TOUT PASSAGE PAR LA SOUS-ROUTINE FUN

CC

IMPLICIT REAL\*8(A-H,O-Z)  
 DIMENSION S(45,10),GY(35),GX(35),V(45),F(45),X(1),H(10),E(45),  
 CSUP(10,10),SA(10)  
 INTEGER P,PI,PU,PA

INITIALISATIONS

CALL FUN(GY,X)  
 CALL SUM(GY,N,SUM2)  
 KAUNT=0  
 KPI=0  
 N1=N+1  
 NP=N+P  
 PA=P+1

IMPRESSIONS INITIALES

PRINT 9012  
 PRINT 301  
 PRINT 422,(X(I),I=1,P)



```

101      PRINT 741
102      PRINT 9003,SUM2
103      PRINT 741
104 C
105 40    KAUNT=KAUNT+1
106      KRI=KRI+1
107 C
108 C      CALCUL DE LA MATRICE DES GRADIENTS
109 C
110      CALL GRAD(S,X)
111      SUM1=SUM2
112      DO 137 I=1,N
113 137    GX(I)=GY(I)
114      DO 920 I=1,N
115 920    F(I)=-GX(I)
116 C
117 C      FACTORISATION DE HOUSEHOLDER DU SYSTEME S*H=F AFIN DE L'AMENER SOUS LA
118 C      FORME S'*H=F' OU S' EST UNE MATRICE TRIANGULAIRE SUPERIEURE DE
119 C      DIMENSIONS (N,P)
120 C
121      DO 904 KI=1,P
122      KI1=KI+1
123      T=0.
124      DO 905 I=KI,N
125 905    T=T+S(I,KI)*#2
126      D=SQRT(T)
127      IF((D-DABS(S(KI,KI)))>.LE.0.0000001) GOTO 904
128      D=-SIGN(D,S(KI,KI))
129      Y1=D*(D-S(KI,KI))
130      DO 906 J=KI,N
131 906    V(J)=S(J,KI)
132      V(KI)=V(KI)-D
133      S(KI,KI)=D
134      IF(KI.EQ.P) GOTO 947
135      DO 907 J=KI1,P
136      DO 908 K=KI,N
137 908    E(K)=S(K,J)
138      DO 907 I=KI,N
139      R=0.
140      DO 909 K=KI,N
141 909    R=R-V(I)*V(K)+E(K)/Y1
142      R=R+S(I,J)
143 907    S(I,J)=R
144 947    DO 910 I=KI,N
145 910    E(I)=F(I)
146      DO 911 I=KI,N
147      R=0.
148      DO 912 K=KI,N
149 912    R=R-V(I)*V(K)+E(K)/Y1
150      R=R+F(I)

```

```

151 911 F(I)=R
152 904 CONTINUE
153 C
154 DO 960 I=1,P
155 DO 961 J=1,P
156 961 SUP(I,J)=S(I,J)
157 960 SA(I)=F(I)
158 KINT=J
159 45 KINT=KINT+1
160 KANT=KANT+1
161 C
162 C CONSTRUCTION DE LA MATRICE AUGMENTEE S'
163 C GAM*I
164 C
165 DO 901 I=N1,NP
166 DO 901 J=1,P
167 S(I-N,J)=SUP(I-N,J)
168 IF(I-J-N)902,903,902
169 902 S(I,J)=0.
170 GOTO 901
171 903 S(I,J)=GAM
172 901 CONTINUE
173 DO 962 J=1,P
174 962 F(J)=SA(J)
175 DO 921 I=N1,NP
176 921 F(I)=0.
177 C
178 C S' F'
179 C FACTORISATION DE HOUSEHOLDER DU SYSTEME *H= AFIN DE L'AMENER SOUS LA
180 C GAM*I C
181 C FORME S'*H=F' OU S' EST UNE MATRICE TRIANGULAIRE SUPERIEURE DE
182 C DIMENSIONS (N+P,P)
183 C
184 DO 931 KI=1,P
185 KI1=KI+1
186 NKI=N+KI
187 T=S(KI,KI)*S(KI,KI)
188 DO 932 I=N1,NKI
189 932 T=T+S(I,KI)*S(I,KI)
190 D=SQRT(T)
191 IF((D-DABS(S(KI,KI)))>.LE.0.0000001) GOTO 931
192 C=-SIGN(D,S(KI,KI))
193 Y1=D*(D-S(KI,KI))
194 DO 933 I=N1,NKI
195 933 V(I)=S(I,KI)
196 V(KI)=S(KI,KI)-D
197 S(KI,KI)=D
198 IF(I.EQ.P) GOTO 950
199 DO 935 J=KI1,P
200 DO 936 K=N1,NKI

```



```

201 936 E(K)=S(K,J)
202 E(KI)=S(KI,J)
203 R=0.
204 DO 941 K=N1,NKI
205 941 R=R-V(KI)*V(K)*E(K)/Y1
206 S(KI,J)=R+S(KI,J)-V(KI)*V(KI)*E(KI)/Y1
207 DO 935 K=N1,NKI
208 R=0.
209 DO 937 I=N1,NKI
210 937 R=R-V(K)*V(I)*E(I)/Y1
211 R=R+S(K,J)
212 935 S(K,J)=R-V(K)*V(KI)*E(KI)/Y1
213 950 DO 939 I=N1,NKI
214 939 E(I)=F(I)
215 E(KI)=F(KI)
216 R=0.
217 DO 943 K=N1,NKI
218 943 R=R-V(KI)*V(K)*E(K)/Y1
219 F(KI)=R+F(KI)-V(KI)*V(KI)*E(KI)/Y1
220 DO 940 J=N1,NKI
221 R=0.
222 DO 952 K=N1,NKI
223 952 R=R-V(J)*V(K)*E(K)/Y1
224 940 F(J)=R+F(J)-V(J)*V(KI)*E(KI)/Y1
225 931 CONTINUE
226 C
227 C DETERMINATION DE H TEL QUE S''*H=F''
228 C
229 PU=P-1
230 H(P)=F(P)/S(P,P)
231 DO 913 I=1,PU
232 R=0.
233 IR=P-I+1
234 DO 914 J=IR,P
235 914 R=F+H(J)*S(IR-1,J)
236 913 H(P-I)=(F(P-I)-R)/S(P-I,P-I)
237 DO 132 I=1,P
238 132 E(I)=X(I)+H(I)
239 C
240 C CALCUL DES QUANTITES NECESSAIRES POUR PASSER A L'ITERATION SUIVANTE
241 C OU POUR SORTIR
242 C
243 SUM3=0.
244 DO 915 I=PA,NP
245 915 SUM3=SUM3+F(I)*A2
246 CALL FUN(GY,E)
247 CALL SUM(GY,N,SUM2)
248 TOTG=SUM1-SUM2
249 SUM3=SUM1-SUM3
250

```

```

251 C    TEST DE SORTIE
252
253      IF(SUM3.LE.EPS) GOTO 50
254 C
255 C    IMPRESSIONS EVENTUELLES
256 C
257      IF(IPRI.LE.0) GOTO 601
258      IF(IPRI.NE.1) GOTO 604
259      PRINT 9013,KAUNT
260      PRINT 9014
261      PRINT 741
262      PRINT 9005,GAM
263      PRINT 741
264      GOTO 605
265 604   IF(KRI.LT.IPRI) GOTO 601
266       PRINT 9013,KAUNT
267       PRINT 9014
268       PRINT 741
269 605   PRINT 9001
270       PRINT 422,(X(I),I=1,P)
271       PRINT 741
272       PRINT 9003,SUM1
273       PRINT 741
274       KRI=0
275 C    ANALYSE DES DIFFERENTS CAS POSSIBLES
276 C
277 601   IF(SIGMA.GT.(TOTG/(2*SUM3))) GOTO 135
278       IF(KINT.EQ.1.AND.(GAM*DECR).GE.0.01) GAM=GAM*DECR
279       DO25I=1,P
280 25     X(I)=E(I)
281 C
282 C    TEST VERIFIANT SI LE NOMBRE MAXIMAL D'ITERATIONS EST ATTEINT OU NON
283 C
284       IF(KAUNT.LT.NMAX) GOTO 40
285       PRINT 9015,NMAX
286       PRINT 741
287       PRINT 9016
288       PRINT 422,(X(I),I=1,P)
289       PRINT 741
290       PRINT 9003,SUM2
291       RETURN
292 135   GAM=GAM*EXPN
293       IF(GAM.GT.GMAX) GOTO 402
294       GOTO 45
295 C
296 C    IMPRESSIONS FINALES
297 C
298 50     PRINT 9012
299       PRINT 9003,KAUNT
300       PRINT 741

```



```

301      PRINT 9004,KANT
302      PRINT 741
303      PRINT 501
304      PRINT 422,(X(I),I=1,P)
305      PRINT 741
306      PRINT 9003,SUM1
307      PRINT 741
308      IF(IPRI.LT.0) GOTO 90
309      PRINT 9017
310      PRINT 741
311      PRINT 422,(GX(I),I=1,N)
312      PRINT 741
313      PRINT 502
314      PRINT 741
315      CALL GRAD(S,X)
316      DO27I=1,N
317 27    PRINT 422,(S(I,J),J=1,P)
318      GOTO 90
319 402   PRINT 503
320      GOTO 90
321 C
322 301   FORMAT(1X,'VECTEUR INITIAL')
323 422   FORMAT(1X,'D16.8')
324 501   FORMAT(1X,'VECTEUR SOLUTION:')
325 502   FORMAT(1X,'MATRICE DES GRADIENTS:')
326 503   FORMAT(1X,'LE PARAMETRE DEPASSE LE MAXIMUM ADMIS')
327 741   FORMAT(1X, '//')
328 9000  FORMAT(1X,'NOMBRE D ITERATIONS:',2X,I3)
329 9001  FORMAT(1X,'NOUVELLE APPROXIMATION:',//)
330 9003  FORMAT(1X,'LA SOMME DES CARRES VAUT:',1X,D16.8)
331 9004  FORMAT(1X,'NOMBRE D ITERATION INTERIEURES:',1X,I3)
332 9005  FORMAT(1X,'VALEUR DU PARAMETRE DE CONTROLE:',1X,D16.8)
333 9012  FORMAT(1H1)
334 9013  FORMAT(1X,'ITERATION:',2X,I3)
335 9014  FORMAT(1X,'*****')
336 9015  FORMAT(1X,'LE NOMBRE MAXIMAL D ITERATIONS ',1X,I3,'EST DEPASSE')
337 9016  FORMAT(1X,'LA MEILLEURE APPROXIMATION OBTENUE EST :')
338 9017  FORMAT(1X,'LA VALEUR DE LA FONCTION EST :')
339 90    RETURN
340      .END

```



```
1 SUBROUTINE SUM(X,P,SUMQ)
2 C
3 C LA SOUS-ROUTINE SUM CALCULE LE CARRE DE LA NORME EUCLIDIENNE D'UN VECTEUR
4 C
5 IMPLICIT REAL*8(A-H,O-Z)
6 DIMENSION X(1)
7 INTEGER P
8 R=0.
9 DO 1 I=1,P
10 1 R=R+X(I)**2
11 SUMQ=R
12 RETURN
13 END
```